

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2006, 20:25:27 ; Search time 180 Seconds
(without alignments)
1883.740 Million cell updates/sec

Title: US-10-649-591-3

Perfect score: 3826

Sequence: 1 LQEVHVSKEITIKISASMKM.....CKCRDQWEGPHCENRFLRRP 732

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA Main:*

- 1: /EMC_Celerra_SIDS3/pcodata/2/pubppa/US07_PUBCOMB.pep:*
- 2: /EMC_Celerra_SIDS3/pcodata/2/pubppa/US08_PUBCOMB.pep:*
- 3: /EMC_Celerra_SIDS3/pcodata/2/pubppa/US09_PUBCOMB.pep:*
- 4: /EMC_Celerra_SIDS3/pcodata/2/pubppa/US10A_PUBCOMB.pep:*
- 5: /EMC_Celerra_SIDS3/pcodata/2/pubppa/US10B_PUBCOMB.pep:*
- 6: /EMC_Celerra_SIDS3/pcodata/2/pubppa/US11_PUBCOMB.pep:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3826	100.0	732	4 US-10-229-345-3	Sequence 3, Appl1
2	3826	100.0	732	4 US-10-274-177-3	Sequence 3, Appl1
3	3826	100.0	732	4 US-10-650-112-3	Sequence 3, Appl1
4	3826	100.0	732	5 US-10-649-591-3	Sequence 3, Appl1
5	3826	100.0	732	6 US-11-041-788-3	Sequence 3, Appl1
6	3826	100.0	732	4 US-10-295-027-124	Sequence 124, App
7	3826	100.0	755	4 US-10-173-999-87	Sequence 87, Appl1
8	3826	100.0	755	4 US-10-229-345-14	Sequence 14, Appl1
9	3826	100.0	755	4 US-10-274-177-14	Sequence 14, Appl1
10	3826	100.0	755	4 US-10-650-112-14	Sequence 14, Appl1
11	3826	100.0	755	5 US-10-649-591-14	Sequence 14, Appl1
12	3826	100.0	755	6 US-11-041-788-14	Sequence 14, Appl1
13	3826	99.9	755	4 US-10-295-027-1186	Sequence 1186, Ap
14	3799	99.3	807	3 US-09-930-020A-2	Sequence 2, Appl1
15	3799	99.3	807	4 US-10-295-027-122	Sequence 122, App
16	3799	99.3	807	4 US-10-295-027-764	Sequence 764, App
17	3799	99.3	807	4 US-10-295-027-777	Sequence 777, App
18	3799	99.3	807	4 US-10-295-027-883	Sequence 883, App
19	3799	99.3	807	4 US-10-295-027-1187	Sequence 1187, Ap
20	3799	99.3	807	4 US-10-087-080-23	Sequence 23, Appl1
21	3799	99.3	807	4 US-10-173-999-85	Sequence 85, Appl1
22	3799	99.3	807	5 US-10-702-180-85	Sequence 2, Appl1
23	3387.5	88.5	1160	5 US-10-450-763-47702	Sequence 47702, A
24	3382.5	88.4	826	4 US-10-295-027-120	Sequence 120, App
25	3382.5	88.4	826	4 US-10-173-999-83	Sequence 83, Appl1
26	2568	67.1	488	5 US-10-649-591-21	Sequence 21, Appl1
27	2568	67.1	488	6 US-11-041-788-21	Sequence 21, Appl1

28	2562	67.0	488	4 US-10-650-112-21	Sequence 21, Appl1
29	1580	41.3	405	3 US-09-866-050A-374	Sequence 374, App
30	522.5	13.7	2211	5 US-10-741-600-1324	Sequence 1324, Ap
31	522.5	13.7	2244	5 US-10-741-600-1321	Sequence 1321, Ap
32	522.5	13.7	2411	5 US-10-741-600-1323	Sequence 1323, Ap
33	522.5	13.7	2977	5 US-10-741-600-1320	Sequence 1320, Ap
34	522.5	13.7	3151	5 US-10-486-678-16	Sequence 16, Appl1
35	522.5	13.7	3173	3 US-09-918-715-218	Sequence 218, App
36	522.5	13.7	3173	4 US-10-474-794-218	Sequence 218, App
37	522.5	13.7	3173	5 US-10-979-159-218	Sequence 218, App
38	522.5	13.7	3176	4 US-10-372-683-20	Sequence 20, Appl1
39	522.5	13.7	3176	4 US-10-734-564-91	Sequence 91, Appl1
40	522.5	13.7	3176	5 US-10-723-860-1065	Sequence 1065, Ap
41	522.5	13.7	3176	5 US-10-486-678-15	Sequence 15, Appl1
42	522.5	13.7	3176	5 US-10-852-335A-160	Sequence 160, App
43	522.5	13.7	3176	5 US-10-287-436A-503	Sequence 503, App
44	522.5	13.7	3176	5 US-10-287-436A-1196	Sequence 1196, App
45	522.5	13.7	3177	5 US-10-741-600-1319	Sequence 1319, Ap

ALIGNMENTS

RESULT 1
US-10-229-345-3
; Sequence 3, Application US//10229345
; Publication No. US20040038220A1
; GENERAL INFORMATION:
; APPLICANT: MARKOWITZ, Sanford D.
; TITLE OF INVENTION: METHODS FOR CATEGORIZING PATIENTS
; FILE REFERENCE: CMRU-P01-003
; CURRENT APPLICATION NUMBER: US/10/229,345
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 732
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-229-345-3

Query Match 100.0%; Score 3826; DB 4; Length 732;

Best Local Similarity 100.0%; Prod. No. 0;

Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	LQEVHVSKEITIKISASMKMCSAAVDIMFLDGSNSVKGKSFERSKHFAITVCGDLDI	60
DB	1	LQEVHVSKEITIKISASMKMCSAAVDIMFLDGSNSVKGKSFERSKHFAITVCGDLDI	60
QY	61	SPERVVGAFOFSSTPHLEFPPLDSFSTQOEVARIKRMVFKGRTETLALKYLHRGUP	120
DB	61	SPERVVGAFOFSSTPHLEFPPLDSFSTQOEVARIKRMVFKGRTETLALKYLHRGUP	120
QY	121	GGNMSVPOILLIVTDGKSQGVVALPSKQIKERGVTFAVGVPRPWEIHALASEPQQ	180
DB	121	GGNMSVPOILLIVTDGKSQGVVALPSKQIKERGVTFAVGVPRPWEIHALASEPQQ	180
QY	121	GGNMSVPOILLIVTDGKSQGVVALPSKQIKERGVTFAVGVPRPWEIHALASEPQQ	180
DB	121	GGNMSVPOILLIVTDGKSQGVVALPSKQIKERGVTFAVGVPRPWEIHALASEPQQ	180
QY	181	HVLLAQVEDATNGLFSTSSNAICSAATPDCRVENHPEHNTLEWRFAGNAPCWRKS	240
DB	181	HVLLAQVEDATNGLFSTSSNAICSAATPDCRVENHPEHNTLEWRFAGNAPCWRKS	240
QY	181	HVLLAQVEDATNGLFSTSSNAICSAATPDCRVENHPEHNTLEWRFAGNAPCWRKS	240
DB	181	HVLLAQVEDATNGLFSTSSNAICSAATPDCRVENHPEHNTLEWRFAGNAPCWRKS	240
QY	241	RRTLAALAHCPYSKRVFLTHPATCYRTTCGPGCDSPQCONGTCVSEGDLGYCCLCP	300
DB	241	RRTLAALAHCPYSKRVFLTHPATCYRTTCGPGCDSPQCONGTCVSEGDLGYCCLCP	300
QY	241	RRTLAALAHCPYSKRVFLTHPATCYRTTCGPGCDSPQCONGTCVSEGDLGYCCLCP	300
DB	241	RRTLAALAHCPYSKRVFLTHPATCYRTTCGPGCDSPQCONGTCVSEGDLGYCCLCP	300
QY	301	LAFGGANCAKLSIECRVDLLFLDSSAGTTLDGLFRAKVPKRVRAVLSEDSRAVYG	360
DB	301	LAFGGANCAKLSIECRVDLLFLDSSAGTTLDGLFRAKVPKRVRAVLSEDSRAVYG	360
QY	301	LAFGGANCAKLSIECRVDLLFLDSSAGTTLDGLFRAKVPKRVRAVLSEDSRAVYG	360
DB	301	LAFGGANCAKLSIECRVDLLFLDSSAGTTLDGLFRAKVPKRVRAVLSEDSRAVYG	360
QY	361	VATYSBELLVAVPVGQYQVDPDLVWLDGIIPRGGTTLGSALROAABGFGSATTTGDD	420
DB	361	VATYSBELLVAVPVGQYQVDPDLVWLDGIIPRGGTTLGSALROAABGFGSATTTGDD	420
QY	361	VATYSBELLVAVPVGQYQVDPDLVWLDGIIPRGGTTLGSALROAABGFGSATTTGDD	420
DB	361	VATYSBELLVAVPVGQYQVDPDLVWLDGIIPRGGTTLGSALROAABGFGSATTTGDD	420

QY 421 RPRRVVVLTTSHSEDEVAGPARHARARELLLLGVSEAVRALEETTSGBKVMYSDP 480
DB 421 RPRRVVVLTTSHSESEBEVAGPARHARARELLLLGVSEAVRALEETTSGBKVMYSDP 480
QY 481 QDLFNOIPELQGLKCSRQRCGRTOALDLVFMLDTSASVGPENFAQMOQSFVRSALQFV 540
DB 481 QDLFNOIPELQGLKCSRQRCGRTOALDLVFMLDTSASVGPENFAQMOQSFVRSALQFV 540
QY 541 NPDVTQGLVTVGSQVQTAFGDLTKPTRAAMLRAISQAPYLGVSAGTALLHIYDKVMT 600
DB 541 NPDVTQGLVTVGSQVQTAFGDLTKPTRAAMLRAISQAPYLGVSAGTALLHIYDKVMT 600
QY 601 VORGARPGVKAVVVLTVGRGAEDAAVPAQKLRNNGISVLYVGVGPVLSGLRRLAGPRD 660
DB 601 VORGARPGVKAVVVLTVGRGAEDAAVPAQKLRNNGISVLYVGVGPVLSGLRRLAGPRD 660
QY 661 SLIHVAAYADLRHYDVLIEMLCGEAKOPVNLCKPSCMNEGSCVYLQNGSYRCKCRDGE 720
DB 661 SLIHVAAYADLRHYDVLIEMLCGEAKOPVNLCKPSCMNEGSCVYLQNGSYRCKCRDGE 720
QY 721 GPHCENRFLRRP 732
DB 721 GPHCENRFLRRP 732

RESULT 2
US-10-274-177-3
; Sequence 3, Application US/10274177
; Publication No. US20040038225A1
; GENERAL INFORMATION:
; APPLICANT: MARKOWITZ, Sanford D.
; TITLE OF INVENTION: METHODS FOR CATEGORIZING PATIENTS
; FILE REFERENCE: CMRU-P01-003
; CURRENT APPLICATION NUMBER: US/10/274,177
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: US/10/229,345
; PRIOR FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; LENGTH: 732
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-274-177-3

Query Match 100.0%; Score 3826; DB 4; Length 732;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LOEVHVSKEITIGKISAASKMMWCSAAVDIMFLDGSNSVGKGSFERSKHFAITVCDGLDI 60
DB 1 LOEVHVSKEITIGKISAASKMMWCSAAVDIMFLDGSNSVGKGSFERSKHFAITVCDGLDI 60
QY 61 SPERRVVGAFQFSSSTPHLEPPLDSFSTQOEVKARIKRMVFKGRTETELALKYLLHRGUP 120
DB 61 SPERRVVGAFQFSSSTPHLEPPLDSFSTQOEVKARIKRMVFKGRTETELALKYLLHRGUP 120
QY 121 GGRNNAVPOILLIIVTDGSGQDVALLPSKQKERGTVFAVGVRFPMEBELHALLASEPRQ 180
DB 121 GGRNNAVPOILLIIVTDGSGQDVALLPSKQKERGTVFAVGVRFPMEBELHALLASEPRQ 180
QY 181 HVLAEQVEDATNGLFSTLSSSAICSSATPDCRVEAHPCERHRTLEWVREFAAGNAPCWRGS 240
DB 181 HVLAEQVEDATNGLFSTLSSSAICSSATPDCRVEAHPCERHRTLEWVREFAAGNAPCWRGS 240
QY 241 RRLTAVLAHCPYSSKRVFLTHPATCYRTTCGPGCDSDPCQNGTGVBEGLDGYQCLCP 300
DB 241 RRLTAVLAHCPYSSKRVFLTHPATCYRTTCGPGCDSDPCQNGTGVBEGLDGYQCLCP 300
QY 301 LAFGEGANCAKLKLSLCRVYDLFLDLSAGTTLDGFLRAKVFYKRFVRAVLSBDSARVG 360
DB 301 LAFGEGANCAKLKLSLCRVYDLFLDLSAGTTLDGFLRAKVFYKRFVRAVLSBDSARVG 360

QY 361 VATYSRELLVAVPGEYQDVPDLWMSLDGI PRGGEPTLTGSAALROAERFGSATRTGOD 420
DB 361 VATYSRELLVAVPGEYQDVPDLWMSLDGI PRGGEPTLTGSAALROAERFGSATRTGOD 420
QY 421 RPRRVVVLTTSHSEBEVAGPARHARARELLLLGVSEAVRALEETTSGBKVMYSDP 480
DB 421 RPRRVVVLTTSHSEBEVAGPARHARARELLLLGVSEAVRALEETTSGBKVMYSDP 480
QY 481 QDLFNOIPELQGLKCSRQRCGRTOALDLVFMLDTSASVGPENFAQMOQSFVRSALQFV 540
DB 481 QDLFNOIPELQGLKCSRQRCGRTOALDLVFMLDTSASVGPENFAQMOQSFVRSALQFV 540
QY 541 NPDVTQGLVTVGSQVQTAFGDLTKPTRAAMLRAISQAPYLGVSAGTALLHIYDKVMT 600
DB 541 NPDVTQGLVTVGSQVQTAFGDLTKPTRAAMLRAISQAPYLGVSAGTALLHIYDKVMT 600
QY 601 VORGARPGVKAVVVLTVGRGAEDAAVPAQKLRNNGISVLYVGVGPVLSGLRRLAGPRD 660
DB 601 VORGARPGVKAVVVLTVGRGAEDAAVPAQKLRNNGISVLYVGVGPVLSGLRRLAGPRD 660
QY 661 SLIHVAAYADLRHYDVLIEMLCGEAKOPVNLCKPSCMNEGSCVYLQNGSYRCKCRDGE 720
DB 661 SLIHVAAYADLRHYDVLIEMLCGEAKOPVNLCKPSCMNEGSCVYLQNGSYRCKCRDGE 720
QY 721 GPHCENRFLRRP 732
DB 721 GPHCENRFLRRP 732

RESULT 3
US-10-650-112-3
; Sequence 3, Application US/10650112
; Publication No. US20040110712A1
; GENERAL INFORMATION:
; APPLICANT: MARKOWITZ, Sanford D.
; TITLE OF INVENTION: METHODS FOR TREATING PATIENTS AND IDENTIFYING THERAPEUTICS
; FILE REFERENCE: CMRU-P01-044
; CURRENT APPLICATION NUMBER: US/10/650,112
; CURRENT FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 10/274,177
; PRIOR FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 10/229,245
; PRIOR FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 60/406,296
; PRIOR FILING DATE: 2002-08-27
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 732
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-650-112-3

Query Match 100.0%; Score 3826; DB 4; Length 732;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LOEVHVSKEITIGKISAASKMMWCSAAVDIMFLDGSNSVGKGSFERSKHFAITVCDGLDI 60
DB 1 LOEVHVSKEITIGKISAASKMMWCSAAVDIMFLDGSNSVGKGSFERSKHFAITVCDGLDI 60
QY 61 SPERRVVGAFQFSSSTPHLEPPLDSFSTQOEVKARIKRMVFKGRTETELALKYLLHRGUP 120
DB 61 SPERRVVGAFQFSSSTPHLEPPLDSFSTQOEVKARIKRMVFKGRTETELALKYLLHRGUP 120
QY 121 GGRNNAVPOILLIIVTDGSGQDVALLPSKQKERGTVFAVGVRFPMEBELHALLASEPRQ 180
DB 121 GGRNNAVPOILLIIVTDGSGQDVALLPSKQKERGTVFAVGVRFPMEBELHALLASEPRQ 180
QY 181 HVLAEQVEDATNGLFSTLSSSAICSSATPDCRVEAHPCERHRTLEWVREFAAGNAPCWRGS 240
DB 181 HVLAEQVEDATNGLFSTLSSSAICSSATPDCRVEAHPCERHRTLEWVREFAAGNAPCWRGS 240

```

Qy 241 RRTLAVLAHCPFYMKRVFLTHPATCYRTTCPCPCDSOPCONGGTCVPEGLDGYOCLCP 300
| | | | |
Db 241 RRTLAVLAHCPFYMKRVFLTHPATCYRTTCPCPCDSOPCONGGTCVPEGLDGYOCLCP 300
Qy 301 LAFGBGANCKLKSIECRVDLLFLDSSAGTTLDGFLRAKVFVKRFVRAVLSEDSBARVG 360
| | | | |
Db 301 LAFGBGANCKLKSIECRVDLLFLDSSAGTTLDGFLRAKVFVKRFVRAVLSEDSBARVG 360
Qy 361 VATTYRELLVAVPVGEYQDPDLVMSLDGI PRFGPTLTGSLARQAEFGSATTGDD 420
| | | | |
Db 361 VATTYRELLVAVPVGEYQDPDLVMSLDGI PRFGPTLTGSLARQAEFGSATTGDD 420
Qy 421 RRRRVVLLTSHSEDEBVGAPARHARARELLLGVSBAVRABEITGSPKHWVYSDP 480
| | | | |
Db 421 RRRRVVLLTSHSEDEBVGAPARHARARELLLGVSBAVRABEITGSPKHWVYSDP 480
Qy 481 QDLFNOIPELOKLCGRORPGCRTOALDLVFMLDTSASVGPENFAQMSFVRSALQFEV 540
| | | | |
Db 481 QDLFNOIPELOKLCGRORPGCRTOALDLVFMLDTSASVGPENFAQMSFVRSALQFEV 540
Qy 541 NPDVTOVGLVYVGSQVQTAFGDLTKPTRAAMLRAISQAPYLGVSAGTALHITDKWMT 600
| | | | |
Db 541 NPDVTOVGLVYVGSQVQTAFGDLTKPTRAAMLRAISQAPYLGVSAGTALHITDKWMT 600
Qy 601 VORGARPGVPAVVVLTGGRGAEDAAVPAQKLRNNGISVLYVGVGVLSBGLRRLAGPRD 660
| | | | |
Db 601 VORGARPGVPAVVVLTGGRGAEDAAVPAQKLRNNGISVLYVGVGVLSBGLRRLAGPRD 660
Qy 661 SLIHVAAYADLRHYQDVLIEMLGCEBAKQPVNLCKPSPCMNEGSCVLYONGSYRCKCDGWE 720
| | | | |
Db 661 SLIHVAAYADLRHYQDVLIEMLGCEBAKQPVNLCKPSPCMNEGSCVLYONGSYRCKCDGWE 720
Qy 721 GPHCENRFLRRP 732
| | | | |
Db 721 GPHCENRFLRRP 732

RESULT 4
US-10-649-591-3
; Sequence 3, Application US/10649591
; Publication No. US20060035237A1
; GENERAL INFORMATION:
; APPLICANT: Markowitz, Sanford D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; FILE REFERENCE: CMRU-P03-003
; CURRENT APPLICATION NUMBER: US/10/649,591
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: US 10/229,345
; PRIOR FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: US 10/274,177
; PRIOR FILING DATE: 2002-10-18
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 732
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-649-591-3

Query Match 100.0%; Score 3826; DB 5; Length 732;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Db 121 GGRNASVPGILLITVDKSGQGDVALPSKQJKEGVTFAVGVRPRMEELHALASEPRGQ 180
| | | | |
Qy 181 HTYLAEOVEDATNGLPSTISSAICSSATPDDRVNAHPCEHHTLBMVRFPAANAPCMWRS 240
| | | | |
Db 181 HTYLAEOVEDATNGLPSTISSAICSSATPDDRVNAHPCEHHTLBMVRFPAANAPCMWRS 240
Qy 241 RRTLAVLAHCPFYMKRVFLTHPATCYRTTCPCPCDSOPCONGGTCVPEGLDGYOCLCP 300
| | | | |
Db 241 RRTLAVLAHCPFYMKRVFLTHPATCYRTTCPCPCDSOPCONGGTCVPEGLDGYOCLCP 300
Qy 301 LAFGBGANCKLKSIECRVDLLFLDSSAGTTLDGFLRAKVFVKRFVRAVLSEDSBARVG 360
| | | | |
Db 301 LAFGBGANCKLKSIECRVDLLFLDSSAGTTLDGFLRAKVFVKRFVRAVLSEDSBARVG 360
Qy 361 VATTYRELLVAVPVGEYQDPDLVMSLDGI PRFGPTLTGSLARQAEFGSATTGDD 420
| | | | |
Db 361 VATTYRELLVAVPVGEYQDPDLVMSLDGI PRFGPTLTGSLARQAEFGSATTGDD 420
Qy 421 RRRRVVLLTSHSEDEBVGAPARHARARELLLGVSBAVRABEITGSPKHWVYSDP 480
| | | | |
Db 421 RRRRVVLLTSHSEDEBVGAPARHARARELLLGVSBAVRABEITGSPKHWVYSDP 480
Qy 481 QDLFNOIPELOKLCGRORPGCRTOALDLVFMLDTSASVGPENFAQMSFVRSALQFEV 540
| | | | |
Db 481 QDLFNOIPELOKLCGRORPGCRTOALDLVFMLDTSASVGPENFAQMSFVRSALQFEV 540
Qy 541 NPDVTOVGLVYVGSQVQTAFGDLTKPTRAAMLRAISQAPYLGVSAGTALHITDKWMT 600
| | | | |
Db 541 NPDVTOVGLVYVGSQVQTAFGDLTKPTRAAMLRAISQAPYLGVSAGTALHITDKWMT 600
Qy 601 VORGARPGVPAVVVLTGGRGAEDAAVPAQKLRNNGISVLYVGVGVLSBGLRRLAGPRD 660
| | | | |
Db 601 VORGARPGVPAVVVLTGGRGAEDAAVPAQKLRNNGISVLYVGVGVLSBGLRRLAGPRD 660
Qy 661 SLIHVAAYADLRHYQDVLIEMLGCEBAKQPVNLCKPSPCMNEGSCVLYONGSYRCKCDGWE 720
| | | | |
Db 661 SLIHVAAYADLRHYQDVLIEMLGCEBAKQPVNLCKPSPCMNEGSCVLYONGSYRCKCDGWE 720
Qy 721 GPHCENRFLRRP 732
| | | | |
Db 721 GPHCENRFLRRP 732

RESULT 5
US-11-041-788-3
; Sequence 3, Application US/11041788
; Publication No. US2005023353A1
; GENERAL INFORMATION:
; APPLICANT: Markowitz, Sanford D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; FILE REFERENCE: CMRU-P04-003
; CURRENT APPLICATION NUMBER: US/11/041,788
; PRIOR FILING DATE: 2005-01-21
; PRIOR APPLICATION NUMBER: US 10/649,591
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: US 10/274,177
; PRIOR FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: US 10/229,345
; PRIOR FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: US 10/087,080
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 60/284,555
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: US 60/281,149
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: US 60/272,206
; PRIOR FILING DATE: 2001-02-27
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 732

```

[illegible]

	APPLICANT:	Murray, Richard	
	APPLICANT:	Watson, Susan R.	
	APPLICANT:	Eos Biotechnology, Inc.	
	TITLE OF INVENTION:	Methods of diagnosis of Cancer, Compositions and	
	TITLE OF INVENTION:	Methods of Screening for Modulators of Cancer	
	FILE REFERENCE:	018501-012500US	
	CURRENT APPLICATION NUMBER:	US/10/295, 027	
	CURRENT FILING DATE:	2002-11-13	
	PRIOR APPLICATION NUMBER:	US 09/663, 733	
	PRIOR FILING DATE:	2000-09-15	
	PRIOR APPLICATION NUMBER:	US 60/350, 666	
	PRIOR FILING DATE:	2001-11-13	
	PRIOR APPLICATION NUMBER:	US 60/335, 394	
	PRIOR FILING DATE:	2001-11-15	
	PRIOR APPLICATION NUMBER:	US 60/332, 464	
	PRIOR FILING DATE:	2001-11-21	
	PRIOR APPLICATION NUMBER:	US 60/334, 393	
	PRIOR FILING DATE:	2001-11-29	
	PRIOR APPLICATION NUMBER:	US 60/340, 376	
	PRIOR FILING DATE:	2001-12-14	
	PRIOR APPLICATION NUMBER:	US 60/347, 211	
	PRIOR FILING DATE:	2002-01-08	
	PRIOR APPLICATION NUMBER:	US 60/347, 349	
	PRIOR FILING DATE:	2002-01-10	
	PRIOR APPLICATION NUMBER:	US 60/355, 250	
	PRIOR FILING DATE:	2002-02-08	
	PRIOR APPLICATION NUMBER:	US 60/356, 714	
	PRIOR FILING DATE:	2002-02-13	
	Remaining Prior Application data removed - See File Wrapper or PALM.		
	NUMBER OF SEQ ID NOS:	1386	
	SOFTWARE:	Patentin Ver. 2.1	
	SEQ ID NO:	124	
	LENGTH:	755	
	TYPE:	PRT	
	ORGANISM:	Homo sapiens	
	US-10-295-027-124		
Query Match	100.0%;	Score 3826;	DB 4; Length 755;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches	732; Conservative	0; Mismatches	0; Indels 0; Gaps 0
QY	1	LOEVVVSXETIGKISAAISKMMWCSAAVIMFLDLSNSVGKSPRSKHFAITVCDDGLI	60
DB	24	LOEVAHVSXETIGKISAAISKMMWCSAAVIMFLDLSNSVGKSPRSKHFAITVCDDGLI	83
QY	61	SPEVRVGAFOFSSSTPHLEPFDSESTOEYKARIKMVPFKGRTELETALRYLLHRGUP	120
DB	84	SPEVRVGAFOFSSSTPHLEPFDSESTOEYKARIKMVPFKGRTELETALRYLLHRGUP	143
QY	121	GGRNAAVPOILLIITYDGKSOGDVALLPSQLKERGTVAAGVFRPRWEELHALASEPRQ	180
DB	144	GGRNAAVPOILLIITYDGKSOGDVALLPSQLKERGTVAAGVFRPRWEELHALASEPRQ	203
QY	181	HVLAEOVEDATNGLFSTLISSAICSSATPCQRVAHPCEHRTLHMVEFPAGNAFCMRGS	240
DB	204	HVLAEOVEDATNGLFSTLISSAICSSATPCQRVAHPCEHRTLHMVEFPAGNAFCMRGS	263
QY	241	RRTIAVIAAHCPFYSMKKVFILTHPATCYRITCCPGCDSPQCQNGTCVPEGIDGYQCLCP	300
DB	264	RRTIAVIAAHCPFYSMKKVFILTHPATCYRITCCPGCDSPQCQNGTCVPEGIDGYQCLCP	323
QY	301	LAFGEANCALKLSIECRVDLIFLLDSAGTTLDGFPAKYFVKRFVAIVLSEDSRARVG	360
DB	324	LAFGEANCALKLSIECRVDLIFLLDSAGTTLDGFPAKYFVKRFVAIVLSEDSRARVG	383
QY	361	VATYSRELVAVPVGEYODVDPLVMSLDGIFFRGGPPTLTGSALROAERGFSSARTTGO	420
DB	384	VATYSRELVAVPVGEYODVDPLVMSLDGIFFRGGPPTLTGSALROAERGFSSARTTGO	443
QY	421	RPRVVVLITESHSEDEVAGPARARARELLLVGSEAVAELBEITGSPGHVWVSDP	480
DB	444	RPRVVVLITESHSEDEVAGPARARARELLLVGSEAVAELBEITGSPGHVWVSDP	503


```
QY 241 RRTLAVALAACHCPFSYMKRVFLTHPATCYRTTCPCPCDSOPCONGTCVPEGLDGYOCLCP 300
DB 264 RRTLAVALAACHCPFSYMKRVFLTHPATCYRTTCPCPCDSOPCONGTCVPEGLDGYOCLCP 323
QY 301 LAFGEANCALKLSLECRVDLLFLDSSAGTTLDDGFLRAKVFYKRFVRAVLSDBSRAV 360
DB 324 LAFGEANCALKLSLECRVDLLFLDSSAGTTLDDGFLRAKVFYKRFVRAVLSDBSRAV 383
QY 361 VATYSRELLVAAPVGEYQDPDLVWSLDGI.PFRGGPTLTGSALRQAERFGSATRTGOD 420
DB 384 VATYSRELLVAAPVGEYQDPDLVWSLDGI.PFRGGPTLTGSALRQAERFGSATRTGOD 443
QY 421 RPRRVVVLITESHSEDEVAAPARHARARELLLLGVGSEAVRALEBITGSPKXVMVYSDP 480
DB 444 RPRRVVVLITESHSEDEVAAPARHARARELLLLGVGSEAVRALEBITGSPKXVMVYSDP 503
QY 481 QDLFNOIPELOGLKCSROPGRTOALD.VFMDLTSAVGPENPAOMOSFVRSALQFEV 540
DB 504 QDLFNOIPELOGLKCSROPGRTOALD.VFMDLTSAVGPENPAOMOSFVRSALQFEV 563
QY 541 NPDVTQVGLVYVGSQVQTA.FGLDTKPTRAMLRALISQAPYLGGVGSAGTALHIYDKVMT 600
DB 564 NPDVTQVGLVYVGSQVQTA.FGLDTKPTRAMLRALISQAPYLGGVGSAGTALHIYDKVMT 623
QY 601 VORGARPGVPAVVVLTGGRGAEDAAVPAOKLRNNGISVLYVGVGPVLSSEGLRRLAGPRD 660
DB 624 VORGARPGVPAVVVLTGGRGAEDAAVPAOKLRNNGISVLYVGVGPVLSSEGLRRLAGPRD 683
QY 661 SLIHVAAYADLRHYDVLIEMLGGEAKOPVNLCKPSPCMNEGSCV.LONGSYRCKCRDGE 720
DB 684 SLIHVAAYADLRHYDVLIEMLGGEAKOPVNLCKPSPCMNEGSCV.LONGSYRCKCRDGE 743
QY 721 GPHCENRFLRRP 732
DB 744 GPHCENRFLRRP 755

RESULT 9
US-10-274-177-14
; Sequence 14, Application US/10274177
; Publication No. US20040038225A1
; GENERAL INFORMATION:
; APPLICANT: MARKOWITZ, Sanford D.
; TITLE OF INVENTION: METHODS FOR CATEGORIZING PATIENTS
; FILE REFERENCE: CMRU-P01-003
; CURRENT APPLICATION NUMBER: US/10/274,177
; PRIOR FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: US/10/229,345
; PRIOR FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 755
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-274-177-14

Query Match 100.0%; Score 3826; DB 4; Length 755;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 181 HVLAEQVEDATNGLFSTLSSAICSSATPDRCVBAHPCEHRTLBMVREFPAGNAPCMRGS 240
DB 204 HVLAEQVEDATNGLFSTLSSAICSSATPDRCVBAHPCEHRTLBMVREFPAGNAPCMRGS 263
QY 241 RRTLAVALAACHCPFSYMKRVFLTHPATCYRTTCPCPCDSOPCONGTCVPEGLDGYOCLCP 300
DB 264 RRTLAVALAACHCPFSYMKRVFLTHPATCYRTTCPCPCDSOPCONGTCVPEGLDGYOCLCP 323
QY 301 LAFGEANCALKLSLECRVDLLFLDSSAGTTLDDGFLRAKVFYKRFVRAVLSDBSRAV 360
DB 324 LAFGEANCALKLSLECRVDLLFLDSSAGTTLDDGFLRAKVFYKRFVRAVLSDBSRAV 383
QY 361 VATYSRELLVAAPVGEYQDPDLVWSLDGI.PFRGGPTLTGSALRQAERFGSATRTGOD 420
DB 384 VATYSRELLVAAPVGEYQDPDLVWSLDGI.PFRGGPTLTGSALRQAERFGSATRTGOD 443
QY 421 RPRRVVVLITESHSEDEVAAPARHARARELLLLGVGSEAVRALEBITGSPKXVMVYSDP 480
DB 444 RPRRVVVLITESHSEDEVAAPARHARARELLLLGVGSEAVRALEBITGSPKXVMVYSDP 503
QY 481 QDLFNOIPELOGLKCSROPGRTOALD.VFMDLTSAVGPENPAOMOSFVRSALQFEV 540
DB 504 QDLFNOIPELOGLKCSROPGRTOALD.VFMDLTSAVGPENPAOMOSFVRSALQFEV 563
QY 541 NPDVTQVGLVYVGSQVQTA.FGLDTKPTRAMLRALISQAPYLGGVGSAGTALHIYDKVMT 600
DB 564 NPDVTQVGLVYVGSQVQTA.FGLDTKPTRAMLRALISQAPYLGGVGSAGTALHIYDKVMT 623
QY 601 VORGARPGVPAVVVLTGGRGAEDAAVPAOKLRNNGISVLYVGVGPVLSSEGLRRLAGPRD 660
DB 624 VORGARPGVPAVVVLTGGRGAEDAAVPAOKLRNNGISVLYVGVGPVLSSEGLRRLAGPRD 683
QY 661 SLIHVAAYADLRHYDVLIEMLGGEAKOPVNLCKPSPCMNEGSCV.LONGSYRCKCRDGE 720
DB 684 SLIHVAAYADLRHYDVLIEMLGGEAKOPVNLCKPSPCMNEGSCV.LONGSYRCKCRDGE 743
QY 721 GPHCENRFLRRP 732
DB 744 GPHCENRFLRRP 755

RESULT 10
US-10-650-112-14
; Sequence 14, Application US/10650112
; Publication No. US20040110712A1
; GENERAL INFORMATION:
; APPLICANT: MARKOWITZ, Sanford D.
; TITLE OF INVENTION: METHODS FOR TREATING PATIENTS AND IDENTIFYING THERAPEUTICS
; FILE REFERENCE: CMRU-P01-044
; CURRENT APPLICATION NUMBER: US/10/650,112
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 10/274,177
; PRIOR FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 10/229,245
; PRIOR FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 60/406,296
; PRIOR FILING DATE: 2002-08-27
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; LENGTH: 755
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-650-112-14

Query Match 100.0%; Score 3826; DB 4; Length 755;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	61	SPERRVACAFQSSSPHLEFPPLDSFSTQOEVRKIRIMVFKGRTTELAKYLHRLGP	120
Db	84	SPERRVAFQFSSSPHLEFPPLDSFSTQOEVRKIRIMVFKGRTTELAKYLHRLGP	143
Qy	121	GGRAVSVEPQILIIYVDGKSGQDVALSPKOLKERGVTVFAGVVPFPMEEHLALASEBRGQ	180
Db	144	GGRAVSVEPQILIIYVDGKSGQDVALSPKOLKERGVTVFAGVVPFPMEEHLALASEBRGQ	203
Qy	161	HVLLAEQVEDATNGLEPSTLISSSAICSSATPDCVEAHPCENRTLEMYREBAGNAPCWRGS	240
Db	204	HVLLAEQVEDATNGLEPSTLISSSAICSSATPDCVEAHPCENRTLEMYREBAGNAPCWRGS	263
Qy	241	RRTLAVLVAHCEPFYSMKRVFLTHPATCYTTCGPGCDSPQCNQGGTCVPEGLDGYQCLCP	300
Db	264	RRTLAVLVAHCEPFYSMKRVFLTHPATCYTTCGPGCDSPQCNQGGTCVPEGLDGYQCLCP	323
Qy	301	LAFGGEANCALKLSLECRVDLLFLDSSAGTTLDDGFRLAVPFKRVFVATLSBDSRAVYG	360
Db	324	LAFGGEANCALKLSLECRVDLLFLDSSAGTTLDDGFRLAKVFPKRVFVATLSBDSRAVYG	383
Qy	361	VATYSREBLVAAPVGEYQDVPDLVWSLDGI.PFRGCGPTLTGSALRQAAERFGSATRTGQD	420
Db	384	VATYSREBLVAAPVGEYQDVPDLVWSLDGI.PFRGCGPTLTGSALRQAAERFGSATRTGQD	443
Qy	421	RPRRVVVLITBSHSEDEVAQPARHARARELLLIGVSEAVRALEETITGS.PKIVMYTSDP	480
Db	444	RPRRVVVLITBSHSEDEVAQPARHARARELLLIGVSEAVRALEETITGS.PKIVMYTSDP	503
Qy	481	ODLPMQIPELOGLKCSRORPGCRTL.DLVFMDTSASVGPENFAMQSPVRSALQFEV	540
Db	504	ODLPMQIPELOGLKCSRORPGCRTL.DLVFMDTSASVGPENFAMQSPVRSALQFEV	563
Qy	541	NPDVTQVGLVVVGSQVQTAFGLDTKPTRAAMLRAISQAPYLGVSAGTALLHIYDKVMT	600
Db	564	NPDVTQVGLVVVGSQVQTAFGLDTKPTRAAMLRAISQAPYLGVSAGTALLHIYDKVMT	623
Qy	601	VORGARPGVPKAVVVLITGGRGAEDAAVPAQKLNNGISVLVGVGVPVLSGLRRLAGPRD	660
Db	624	VORGARPGVPKAVVVLITGGRGAEDAAVPAQKLNNGISVLVGVGVPVLSGLRRLAGPRD	683
Qy	661	SLIHVAAYADLRYHODVLEWLCGEAKOPVNLCKBSPCMNBSGCVLQNGSYRCKCDGWE	720
Db	684	SLIHVAAYADLRYHODVLEWLCGEAKOPVNLCKBSPCMNBSGCVLQNGSYRCKCDGWE	743
Qy	721	GPHCENRFLRRP 732	
Db	744	GPHCENRFLRRP 755	
RESULT 11			
US-10-649-591-14			
; Sequence 14, Application US/10649591			
; Publication No. US20060035237A1			
; GENERAL INFORMATION:			
; APPLICANT: Markowitz, Sanford D.			
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR			
; FILE REFERENCE: CWRU-P03-003			
; CURRENT APPLICATION NUMBER: US/10/649,591			
; CURRENT FILING DATE: 2003-08-26			
; PRIOR APPLICATION NUMBER: US 10/229,345			
; PRIOR FILING DATE: 2002-08-26			
; PRIOR APPLICATION NUMBER: US 10/274,177			
; PRIOR FILING DATE: 2002-10-18			
; NUMBER OF SEQ ID NOS: 22			
; SOFTWARE: FaelsBo for Windows Version 4.0			
; SEQ ID NO 14			
; LENGTH: 755			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-649-591-14			
Query Match 100.0%; Score 3826; DB 5; Length 755;			

[illegible]

QY 301 LAFGGANALKLSLECRVDLFLDSSAGTTLDGFLRAKVFYKRFVRAVLSEDSBARVG 360
 DB 324 LAFGGANALKLSLECRVDLFLDSSAGTTLDGFLRAKVFYKRFVRAVLSEDSBARVG 383
 QY 361 VATTYRELLVAAPVGEYQDPDLVMSLDGI PFRGGPTLTGSALRQAERGFSGATRTGOD 420
 DB 384 VATTYRELLVAAPVGEYQDPDLVMSLDGI PFRGGPTLTGSALRQAERGFSGATRTGOD 443
 QY 421 RRRRVVLLTSHSEDEBVAGPARHARARELLLGVSSEAVRALEBITGSPKRWVYSDP 480
 DB 444 RRRRVVLLTSHSEDEBVAGPARHARARELLLGVSSEAVRALEBITGSPKRWVYSDP 503
 QY 481 ODLENOIPELOGLCSRRPGCRTOALDVFMLDTSASVGPENFAQMSFVSCALQFEY 540
 DB 504 ODLENOIPELOGLCSRRPGCRTOALDVFMLDTSASVGPENFAQMSFVSCALQFEY 563
 QY 541 NPDVTQGLVVGSGOQTAFGLDTKPTRAAMLRAISQAPYLGVSAGTALLHIYDKWMT 600
 DB 564 NPDVTQGLVVGSGOQTAFGLDTKPTRAAMLRAISQAPYLGVSAGTALLHIYDKWMT 623
 QY 601 VORGARPGVPAKVVVLTGGRGAEDAAVPAOKLRNNGISVLYVGVGVYLSBGLRRLAGPRD 660
 DB 624 VORGARPGVPAKVVVLTGGRGAEDAAVPAOKLRNNGISVLYVGVGVYLSBGLRRLAGPRD 683
 QY 661 SLIHVAAYADLRHYHODVLLEWLCGBAKOPVNLCKSPCMNBSCVYLQNSYRCKCKRDGWE 720
 DB 684 SLIHVAAYADLRHYHODVLLEWLCGBAKOPVNLCKSPCMNBSCVYLQNSYRCKCKRDGWE 743
 QY 721 GPHCENRFLRRP 732
 DB 744 GPHCENRFLRRP 755

RESULT 14
 US-09-930-020A-2
 ; Sequence 2, Application US/0930020A
 ; Publication No. US20030077568A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gish, Kurt C.
 ; APPLICANT: Mack, David H.
 ; APPLICANT: Wilson, Keith B.
 ; APPLICANT: Eos Biotechnology, Inc.
 ; TITLE OF INVENTION: Methods of diagnosis of Colorectal Cancer. Compositions
 ; TITLE OF INVENTION: and Methods of Screening for Colorectal Cancer
 ; TITLE OF INVENTION: Modulators
 ; FILE REFERENCE: 018501-003100US
 ; CURRENT APPLICATION NUMBER: US/09/930,020A
 ; CURRENT FILING DATE: 2001-08-14
 ; PRIOR APPLICATION NUMBER: US 09/663,733
 ; PRIOR FILING DATE: 2000-09-15
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 807
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: CBF9
 US-09-930-020A-2

Query Match 99.3%; Score 3799; DB 3; Length 807;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 727; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LOEVHNSKETIGKISASAKMMSCAAVDIMFLDGSNSVKGSPERSKHPAIVTCGLDI 60
 DB 24 LOEVHNSKETIGKISASAKMMSCAAVDIMFLDGSNSVKGSPERSKHPAIVTCGLDI 83
 QY 61 SPERVAVGAFQSFSTPLLEPPLDSFSTQGVKARIKRMVKGKRTTELAKTLIRGUP 120
 DB 84 SPERVAVGAFQSFSTPLLEPPLDSFSTQGVKARIKRMVKGKRTTELAKTLIRGUP 143
 QY 121 GGNNAVPOQLLIVTDGKSGDVALPSKQKRGVTVFAVGAFPRMBELHALASEPRG 180

DB 144 GGNNAVPOQLLIVTDGKSGDVALPSKQKRGVTVFAVGAFPRMBELHALASEPRG 203
 QY 181 HTLLAEQVEDATNGFLSTSSAIGSATPPDRCVRAHPERBHTLEWVRPAGNAPCMRRS 240
 DB 204 HTLLAEQVEDATNGFLSTSSAIGSATPPDRCVRAHPERBHTLEWVRPAGNAPCMRRS 263
 QY 241 RRTLAVLAHCPFYSGMKRVFLTHPATCYRTTCGPPCDSQPCONGTCVEBGLDGYOCLP 300
 DB 264 RRTLAVLAHCPFYSGMKRVFLTHPATCYRTTCGPPCDSQPCONGTCVEBGLDGYOCLP 323
 QY 301 LAFGGANALKLSLECRVDLFLDSSAGTTLDGFLRAKVFYKRFVRAVLSEDSBARVG 360
 DB 324 LAFGGANALKLSLECRVDLFLDSSAGTTLDGFLRAKVFYKRFVRAVLSEDSBARVG 383
 QY 361 VATTYRELLVAAPVGEYQDPDLVMSLDGI PFRGGPTLTGSALRQAERGFSGATRTGOD 420
 DB 384 VATTYRELLVAAPVGEYQDPDLVMSLDGI PFRGGPTLTGSALRQAERGFSGATRTGOD 443
 QY 421 RRRRVVLLTSHSEDEBVAGPARHARARELLLGVSSEAVRALEBITGSPKRWVYSDP 480
 DB 444 RRRRVVLLTSHSEDEBVAGPARHARARELLLGVSSEAVRALEBITGSPKRWVYSDP 503
 QY 481 ODLENOIPELOGLCSRRPGCRTOALDVFMLDTSASVGPENFAQMSFVSCALQFEY 540
 DB 504 ODLENOIPELOGLCSRRPGCRTOALDVFMLDTSASVGPENFAQMSFVSCALQFEY 563
 QY 541 NPDVTQGLVVGSGOQTAFGLDTKPTRAAMLRAISQAPYLGVSAGTALLHIYDKWMT 600
 DB 564 NPDVTQGLVVGSGOQTAFGLDTKPTRAAMLRAISQAPYLGVSAGTALLHIYDKWMT 623
 QY 601 VORGARPGVPAKVVVLTGGRGAEDAAVPAOKLRNNGISVLYVGVGVYLSBGLRRLAGPRD 660
 DB 624 VORGARPGVPAKVVVLTGGRGAEDAAVPAOKLRNNGISVLYVGVGVYLSBGLRRLAGPRD 683
 QY 661 SLIHVAAYADLRHYHODVLLEWLCGBAKOPVNLCKSPCMNBSCVYLQNSYRCKCKRDGWE 720
 DB 684 SLIHVAAYADLRHYHODVLLEWLCGBAKOPVNLCKSPCMNBSCVYLQNSYRCKCKRDGWE 743
 QY 721 GPHCENR 727
 DB 744 GPHCENR 750

RESULT 15
 US-10-295-027-122
 ; Sequence 122, Application US/10295027
 ; Publication No. US20030232350A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Afari, Daniel
 ; APPLICANT: Aziz, Nacasha
 ; APPLICANT: Ginsberg, Wendy M.
 ; APPLICANT: Gish, Kurt C.
 ; APPLICANT: Glynn, Richard
 ; APPLICANT: Hevez, Peter A.
 ; APPLICANT: Mack, David H.
 ; APPLICANT: Murray, Richard
 ; APPLICANT: Watson, Susan R.
 ; APPLICANT: Eos Biotechnology, Inc.
 ; TITLE OF INVENTION: Methods of diagnosis of Cancer. Compositions and
 ; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
 ; FILE REFERENCE: 018501-012500US
 ; CURRENT APPLICATION NUMBER: US/10/295,027
 ; CURRENT FILING DATE: 2002-11-13
 ; PRIOR APPLICATION NUMBER: US 09/663,733
 ; PRIOR FILING DATE: 2000-09-15
 ; PRIOR APPLICATION NUMBER: US 60/350,666
 ; PRIOR FILING DATE: 2001-11-13
 ; PRIOR APPLICATION NUMBER: US 60/335,394
 ; PRIOR FILING DATE: 2001-11-15
 ; PRIOR APPLICATION NUMBER: US 60/332,464
 ; PRIOR FILING DATE: 2001-11-21
 ; PRIOR APPLICATION NUMBER: US 60/334,393

; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 122
; LENGTH: 807
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-122

Query Match 99.3%; Score 3799; DB 4; Length 807;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 727; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQEVHYSKETIGKISAAKMMCSAVIDMFLDGSNSVGKSPERSKHPAIVCDGLDI 60
DB 24 LQEVHYSKETIGKISAAKMMCSAVIDMFLDGSNSVGKSPERSKHPAIVCDGLDI 83
QY 61 SPERVAVGAFOFSTSTPHLEFPPLDSFSTQDEVKARIKRMVFKGRTETELAKYLHRLP 120
DB 84 SPERVAVGAFOFSTSTPHLEFPPLDSFSTQDEVKARIKRMVFKGRTETELAKYLHRLP 143
QY 121 GGNASVPOLLIIIVTDGKSQGDVALPSKQIKERGVTFAVGVPFPMWELHALASEPRQ 180
DB 144 GGNASVPOLLIIIVTDGKSQGDVALPSKQIKERGVTFAVGVPFPMWELHALASEPRQ 203
QY 181 HVLAAQVEDATNGLFSTLSSSAICSSATPDCEVAHPCHEHRTLEWREBPAGNAPCMRGS 240
DB 204 HVLAAQVEDATNGLFSTLSSSAICSSATPDCEVAHPCHEHRTLEWREBPAGNAPCMRGS 263
QY 241 RRLTAVLAACHCFYSWKRVEFLTHPATCYRTTCGPGCDSDPCQNGTCEPBGIDGYOCLCP 300
DB 264 RRLTAVLAACHCFYSWKRVEFLTHPATCYRTTCGPGCDSDPCQNGTCEPBGIDGYOCLCP 323
QY 301 LAEGSEANCAKLISLECRVDLFLDLSAGTTLDGFLRAKVPYKRFVRAVLSSEDSRAVG 360
DB 324 LAEGSEANCAKLISLECRVDLFLDLSAGTTLDGFLRAKVPYKRFVRAVLSSEDSRAVG 383
QY 361 VATTYSRELLVAVPVGEYQVDPDLVMSLDGIPFRGSPITLGSALROAERFGSATRTGDD 420
DB 384 VATTYSRELLVAVPVGEYQVDPDLVMSLDGIPFRGSPITLGSALROAERFGSATRTGDD 443
QY 421 RPRRVVLLTSHSEDEVAHPARHARARELLLLGVSEBAVRALEBEITGSPKHVMYSDP 480
DB 444 RPRRVVLLTSHSEDEVAHPARHARARELLLLGVSEBAVRALEBEITGSPKHVMYSDP 503
QY 481 QDLFNGIPELQGLCSRRQPGCCTQALDIVFMLDTSASVGPENFAQMGFVRSICALQFEV 540
DB 504 QDLFNGIPELQGLCSRRQPGCCTQALDIVFMLDTSASVGPENFAQMGFVRSICALQFEV 563
QY 541 NPDVVOVGLVYVGSQVOTAFGLDTKPTRAMLRAISQAPYLGGVGSAGTALLHIYDKVMT 600
DB 564 NPDVVOVGLVYVGSQVOTAFGLDTKPTRAMLRAISQAPYLGGVGSAGTALLHIYDKVMT 623
QY 601 VORGARPGVPAKAVVLTGGRGABDAVPAOKLRNNGISVTVGVGPVLSSEGLRRLAGPRD 660
DB 624 VORGARPGVPAKAVVLTGGRGABDAVPAOKLRNNGISVTVGVGPVLSSEGLRRLAGPRD 683
QY 661 SLIHVAAYADLRYHQDVLLEWLCGEAKOPVNLCKPSPCMNBSGCVLQNGSYRCKCRDGMW 720
DB 684 SLIHVAAYADLRYHQDVLLEWLCGEAKOPVNLCKPSPCMNBSGCVLQNGSYRCKCRDGMW 743
QY 721 GPHCENR 727

Db 744 GPHCENR 750
Search completed: June 21, 2006, 20:28:48
Job time : 183 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2006, 20:04:13 ; Search time 201 Seconds

(without alignments)
1665.087 Million cell updates/sec

Title: US-10-649-591-3

Perfect score: 3826

Sequence: 1 LQEVHVKETIGKISASKM.....CKCRDGEGRPCENFLARP 732

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_8:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*
- 9: geneseqp2005s:*
- 10: geneseqp2006s:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3826	100.0	732	8	ADL70216
2	3826	100.0	732	8	ADN59608
3	3826	100.0	755	6	ABUS6623
4	3826	100.0	755	7	ADN80523
5	3826	100.0	755	7	ADN38806
6	3826	100.0	755	8	ADL70227
7	3826	100.0	755	8	ADN59619
8	3826	99.9	755	7	ADN39868
9	3799	99.3	807	5	ABG61891
10	3799	99.3	807	5	ABP54687
11	3799	99.3	807	6	ABUS6622
12	3799	99.3	807	7	ADN80521
13	3799	99.3	807	7	ADN39869
14	3799	99.3	807	7	ADN38804
15	3799	99.3	807	7	ADN39446
16	3799	99.3	807	7	ADN39459
17	3799	99.3	807	8	ADN39565
18	3799	99.3	807	8	ADN70146
19	3540	92.5	725	8	ADN65732
20	3528	92.2	725	8	ADN809513
21	3387.5	88.5	1160	4	ABG17343
22	3382.5	88.4	826	6	ABUS6621
23	3382.5	88.4	826	7	ADN80519

24	3382.5	88.4	826	7	ADN38802	Adn38802	Cancer/an
25	3177.5	83.1	731	8	ADG78854	Human can	
26	2568	67.1	488	8	ADL70234	Colon can	
27	2568	67.1	488	8	ADN59626	Colon neo	
28	1580	41.3	405	3	AAV76095	Murine bk	
29	1580	41.3	405	4	AA56034	Skin cell	
30	1580	41.3	405	5	AB872234	Murine pr	
31	522.5	13.7	2000	9	ADP54237	Human PRO	
32	522.5	13.7	2000	9	ADY14520	PRO POLYP	
33	522.5	13.7	2211	8	ADQ39661	Human myo	
34	522.5	13.7	2244	8	ADQ39658	Human myo	
35	522.5	13.7	2411	8	ADQ39660	Human myo	
36	522.5	13.7	2977	8	ABW82566	Human dia	
37	522.5	13.7	2977	8	ADQ39657	Human myo	
38	522.5	13.7	3151	6	ABR39571	Human alp	
39	522.5	13.7	3173	5	ABR90743	Human tum	
40	522.5	13.7	3173	6	ABU54450	Human tum	
41	522.5	13.7	3176	6	ABR39570	Human alp	
42	522.5	13.7	3176	7	ADP65255	Human alp	
43	522.5	13.7	3176	8	ADQ18247	Human sof	
44	522.5	13.7	3176	8	ADQ29665	Human col	
45	522.5	13.7	3176	8	ADR87616	Human typ	

ALIGNMENTS

RESULT 1
ID ADL70216
ADL70216 standard; protein; 732 AA.
XX
AC ADL70216;
XX
DT 20-MAY-2004 (first entry)
XX
DE Colon cancer marker ColoUp2 secreted protein.
XX
DE ColoUp2; colon cancer; marker; diagnosis; cytostatic; vaccine;
XX
KW gene therapy.
XX
KW
XX
OS Homo sapiens.
XX
PN WO2004018648-A2.
XX
PD 04-MAR-2004.
XX
PF 26-AUG-2003; 2003WO-US027086.
XX
PR 26-AUG-2002; 2002US-00229345.
XX
PR 27-AUG-2002; 2002US-0406296P.
XX
PR 18-OCT-2002; 2002US-00274117.
XX
XX (UYCA-) UNIV CASE WESTERN RESERVE.
XX
PI Markowitz SD;
XX
DR WPI; 2004-226834/21.
XX
DR N-PSDB; ADL70218.
XX
PT Inhibiting the growth or proliferation of a colon neoplasia in a subject,
XX
PT for treating colon cancer, comprises administering to the subject an
XX
PT agent that decreases the amount of a polypeptide present in or produced
XX
PT by the colon neoplasia.
XX
PS Claim 10; SEQ ID NO 3; 125pp; English.
XX
CC The present sequence is the protein sequence of secreted ColoUp2, a
XX
CC molecular marker of colon neoplasia. ColoUp2 was identified by expression
XX
CC microarray profiling of genes showing differential expression in normal
XX
CC colon and metastatic colon cancer. ColoUp2 transcript is essentially
XX
CC undetectable in normal colon epithelial strips, in normal and in colonic
XX
CC muscle, but expression is clearly detectable in premalignant colon
XX
CC adenomas and in 90% of Duke's stage B (early negative colon cancers),

CC Dukes stage C (node positive colon cancer), Dukes stage D (primary colon
 CC cancers with associated metastatic spread) and in colon cancer liver
 CC metastasis. Expression is also detectable in colon cancer cell lines and
 CC in colon cancer xenografts grown in athymic mice. Col01p2 is secreted
 CC from both the basolateral and apical surfaces of intestinal cells, and
 CC can be found in the blood in 2 different forms, a full-length secreted
 CC form and a C-terminal fragment. The molecular marker is useful in the
 CC detection or diagnosis of colon neoplasia and for categorizing the
 CC neoplastic state of a patient. Methods for inhibiting the growth or
 CC proliferation of a colon neoplasia involve the use of short interfering
 CC RNA (siRNA) or antisense oligonucleotides, or an agent that binds to and
 CC antagonises the polypeptide. Molecular markers can also be used to target
 CC chemotherapeutic agents to cells of a colon neoplasia, and to screen and
 CC identify candidate therapeutic agents.

XX
 XX
 SQ Sequence 732 AA;

Query Match 100.0%; Score 3826; DB 8; Length 732;

Best Local Similarity 100.0%; Pred. No. 6.3e-300;

Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQEVHVSKEITIGKISAAKMMCSAAVDIMFLDGSNSVGKSPERSKHFAITVCDGLDI 60
 DB 1 LQEVHVSKEITIGKISAAKMMCSAAVDIMFLDGSNSVGKSPERSKHFAITVCDGLDI 60
 QY 61 SPERVRVGAFOFSSSTPHLEFPDLSFSTQOEYKARIKRWVFKGRTETELAKYLLHRLGP 120
 DB 61 SPERVRVGAFOFSSSTPHLEFPDLSFSTQOEYKARIKRWVFKGRTETELAKYLLHRLGP 120
 QY 121 GGRNASVPOIILITVDGSGQDVALPSKQKRGVTVFAVGVRPRMELHALLASPRQ 180
 DB 121 GGRNASVPOIILITVDGSGQDVALPSKQKRGVTVFAVGVRPRMELHALLASPRQ 180
 QY 181 HYLAEQVEDATNGLFSTLSSSAICSSATPDCRVEAHPCERTLEWREFFAGNAPCWRGS 240
 DB 181 HYLAEQVEDATNGLFSTLSSSAICSSATPDCRVEAHPCERTLEWREFFAGNAPCWRGS 240
 QY 241 RRTLAVALAHCPEYSMKRVFLTHPATCYRTTCPCDSDQPCQNGTCVPEGLDGYQCLCP 300
 DB 241 RRTLAVALAHCPEYSMKRVFLTHPATCYRTTCPCDSDQPCQNGTCVPEGLDGYQCLCP 300
 QY 301 LAFGGANCAKLKLSLECRVDLFLDSSAGTTIDGFLRAKVPYKRVRAVLSEDSRARVG 360
 DB 301 LAFGGANCAKLKLSLECRVDLFLDSSAGTTIDGFLRAKVPYKRVRAVLSEDSRARVG 360
 QY 361 VATYSBELVAVPVGEYQVDPDLVWGLDGLPRFGGTTLGSALRQAERGFSGATRTGQD 420
 DB 361 VATYSBELVAVPVGEYQVDPDLVWGLDGLPRFGGTTLGSALRQAERGFSGATRTGQD 420
 QY 421 RPRRVVVILLTSHSEDEVAAPARHARARELLLVGSEAVRAELEBITGSPKRWVYSDP 480
 DB 421 RPRRVVVILLTSHSEDEVAAPARHARARELLLVGSEAVRAELEBITGSPKRWVYSDP 480
 QY 481 QDLFNOIPELQKLCGRQRPGRCTQALDIVFMLDTSASVGPENFAMQSPVSSCALQFEV 540
 DB 481 QDLFNOIPELQKLCGRQRPGRCTQALDIVFMLDTSASVGPENFAMQSPVSSCALQFEV 540
 QY 541 NPDVTOVGAVVYGSQVQTAFGDLTKPTRAAMLRAISQAPYLGVSAGATLHITDKWMT 600
 DB 541 NPDVTOVGAVVYGSQVQTAFGDLTKPTRAAMLRAISQAPYLGVSAGATLHITDKWMT 600
 QY 601 VORGARPGVPAKVVVLTGGRGAEDAAVPAQKLRNNGISVLYVGVGVLSSEGLRRLAGPRD 660
 DB 601 VORGARPGVPAKVVVLTGGRGAEDAAVPAQKLRNNGISVLYVGVGVLSSEGLRRLAGPRD 660
 QY 661 SLIHVAAYADLRHNOVLLEWLGEAKQPVNLCKPSPCMNBSGCVLQNSYCKCRDGE 720
 DB 661 SLIHVAAYADLRHNOVLLEWLGEAKQPVNLCKPSPCMNBSGCVLQNSYCKCRDGE 720
 QY 721 GPHCENRFLRRP 732
 DB 721 GPHCENRFLRRP 732

RESULT 2
 ADNS9608
 ID ADNS9608 brandard; protein; 732 AA.

XX AC ADNS9608;

XX DT 01-JUL-2004 (first entry)

XX DE Colon neoplasia Col01p2 protein, SEQ ID No 3.

XX KW molecular marker; neoplastic state; colon neoplasia; Col01p1; Col01p2;
 KW colon cancer.

XX OS Unidentified.

XX PN WO2004018647-A2.

XX PD 04-MAR-2004.

XX PF 26-AUG-2003; 2003WO-US027085.

XX PR 26-AUG-2002; 2002US-00229345.

XX PR 18-OCT-2002; 2002US-00274177.

XX PA (UYCA-) UNIV CASE WESTERN RESERVE.

XX PI Markowitz SD;

XX DR WPI; 2004-419536/39.

XX PT Detecting likelihood of subject to have colon neoplasia, involves
 PT obtaining biological sample from subject, detecting one or more of
 PT Col01p1 and Col01p2 polypeptides, presence of polypeptide indicating
 PT colon neoplasia.

XX PS Claim 11; SEQ ID NO 3; 118pp; English.

XX SQ Sequence 732 AA;

XX SQ Sequence 732 AA;

The invention relates to novel molecular markers for categorizing the
 CC neoplastic state of a patient. The invention further comprises a method
 CC for detecting whether a subject is likely to have a colon neoplasia. The
 CC method comprises: obtaining a biological sample from the subject;
 CC detecting one or more polypeptides selected from among: one or more
 CC Col01p1 polypeptides and one or more Col01p2 polypeptides, wherein the
 CC presence of said one or more polypeptides is indicative of colon
 CC neoplasia. The invention further comprises a kit for carrying out the
 CC said method. The methods and compositions are useful for detecting colon
 CC cancer. This sequence represents a Col01p2 protein of the invention.

XX SQ Sequence 732 AA;

Query Match 100.0%; Score 3826; DB 8; Length 732;

Best Local Similarity 100.0%; Pred. No. 6.3e-300;

Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQEVHVSKEITIGKISAAKMMCSAAVDIMFLDGSNSVGKSPERSKHFAITVCDGLDI 60
 DB 1 LQEVHVSKEITIGKISAAKMMCSAAVDIMFLDGSNSVGKSPERSKHFAITVCDGLDI 60
 QY 61 SPERVRVGAFOFSSSTPHLEFPDLSFSTQOEYKARIKRWVFKGRTETELAKYLLHRLGP 120
 DB 61 SPERVRVGAFOFSSSTPHLEFPDLSFSTQOEYKARIKRWVFKGRTETELAKYLLHRLGP 120
 QY 121 GGRNASVPOIILITVDGSGQDVALPSKQKRGVTVFAVGVRPRMELHALLASPRQ 180
 DB 121 GGRNASVPOIILITVDGSGQDVALPSKQKRGVTVFAVGVRPRMELHALLASPRQ 180
 QY 181 HYLAEQVEDATNGLFSTLSSSAICSSATPDCRVEAHPCERTLEWREFFAGNAPCWRGS 240
 DB 181 HYLAEQVEDATNGLFSTLSSSAICSSATPDCRVEAHPCERTLEWREFFAGNAPCWRGS 240
 QY 241 RRTLAVALAHCPEYSMKRVFLTHPATCYRTTCPCDSDQPCQNGTCVPEGLDGYQCLCP 300
 DB 241 RRTLAVALAHCPEYSMKRVFLTHPATCYRTTCPCDSDQPCQNGTCVPEGLDGYQCLCP 300

D	b	241	KRTLA	VLA	ANHC	PF	YMKV	VEL	TH	PAT	CY	TT	C	B	P	D	S	P	Q	N	G	T	C	V	E	B	G	D	V	Y	C	L	R	300																										
Q	y	301	LA	P	G	E	A	N	C	A	L	K	L	S	E	C	R	V	D	L	L	P	L	D	S	A	G	T	T	L	D	F	L	R	A	K	V	P	Y	K	F	V	R	A	V	L	S	E	D	S	A	R	V	G	360					
D	b	301	LA	P	G	E	A	N	C	A	L	K	L	S	E	C	R	V	D	L	L	P	L	D	S	A	G	T	T	L	D	F	L	R	A	K	V	P	Y	K	F	V	R	A	V	L	S	E	D	S	A	R	V	G	360					
Q	y	361	VAT	Y	R	E	L	L	V	A	P	V	S	E	O	P	V	P	L	W	S	L	D	I	P	F	G	G	T	L	G	S	A	L	R	O	A	E	R	G	F	S	A	T	T	G	D	420												
D	b	361	VAT	Y	R	E	L	L	V	A	P	V	S	E	O	P	V	P	L	W	S	L	D	I	P	F	G	G	T	L	G	S	A	L	R	O	A	E	R	G	F	S	A	T	T	G	D	420												
Q	y	421	VAT	Y	R	E	L	L	V	A	P	V	S	E	O	P	V	P	L	W	S	L	D	I	P	F	R	G	G	T	L	G	S	A	L	R	O	A	E	R	G	F	S	A	T	T	G	D	420											
D	b	421	VAT	Y	R	E	L	L	V	A	P	V	S	E	O	P	V	P	L	W	S	L	D	I	P	F	R	G	G	T	L	G	S	A	L	R	O	A	E	R	G	F	S	A	T	T	G	D	420											
Q	y	421	R	P	R	V	V	L	L	T	S	H	S	E	D	E	V	A	G	P	A	R	H	A	R	E	L	L	I	G	V	S	E	A	V	R	A	E	I	T	S	P	K	H	V	M	Y	S	D	480										
D	b	421	R	P	R	V	V	L	L	T	S	H	S	E	D	E	V	A	G	P	A	R	H	A	R	E	L	L	I	G	V	S	E	A	V	R	A	E	I	T	S	P	K	H	V	M	Y	S	D	480										
Q	y	481	O	D	L	F	M	O	I	P	E	L	O	K	L	C	S	R	O	R	P	G	C	T	O	A	L	D	L	V	F	M	L	D	T	S	A	S	G	P	E	N	F	A	O	M	O	S	F	V	R	S	C	A	L	O	F	E	V	540
D	b	481	O	D	L	F	M	O	I	P	E	L	O	K	L	C	S	R	O	R	P	G	C	T	O	A	L	D	L	V	F	M	L	D	T	S	A	S	G	P	E	N	F	A	O	M	O	S	F	V	R	S	C	A	L	O	F	E	V	540
Q	y	541	N	E	D	V	T	O	V	G	L	V	V	G	S	O	V	O	T	A	F	G	L	D	T	K	P	T	R	A	M	L	R	A	I	S	O	A	P	Y	L	G	V	S	A	G	T	L	L	H	I	Y	D	K	V	M	T	600		
D	b	541	N	E	D	V	T	O	V	G	L	V	V	G	S	O	V	O	T	A	F	G	L	D	T	K	P	T	R	A	M	L	R	A	I	S	O	A	P	Y	L	G	V	S	A	G	T	L	L	H	I	Y	D	K	V	M	T	600		
Q	y	601	V	O	R	G	A	R	P	G	V	K	V	V	L	T	G	R	G	A	E	D	A	V	P	A	O	K	L	R	N	N	G	I	S	V	L	V	V	G	V	G	L	S	E	G	L	R	L	A	G	P	R	660						
D	b	601	V	O	R	G	A	R	P	G	V	K	V	V	L	T	G	R	G	A	E	D	A	V	P	A	O	K	L	R	N	N	G	I	S	V	L	V																						

XX	RESULT 3
XX	ABUS6623
ID	ABUS6623 standard; protein; 755 AA.
XX	
AC	ABUS6623;
XX	
DT	02-APR-2003 (first entry)
XX	
DE	Lung cancer-associated polypeptide #216.
XX	
KW	Lung cancer-associated polypeptide; emphysema;
KW	antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW	small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW	chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW	interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX	
OS	Unidentified.
XX	
XX	WO200286443-A2.
PN	
XX	
PD	31-OCT-2002.
XX	
PF	18-APR-2002; 2002MO-US012476.
XX	
XX	
PR	18-APR-2001; 2001US-0284770P.
PR	10-MAY-2001; 2001US-0290492P.
PR	09-NOV-2001; 2001US-0339245P.
PR	13-NOV-2001; 2001US-0350666P.
PR	29-NOV-2001; 2001US-0334370P.
PR	12-APR-2002; 2002US-0372246P.
XX	
XX	
PA	(EOSB-) EOS BIOTECHNOLOGY INC.
XX	
PT	Aziz N, Murray R;
XX	
XX	WPI, 2003-093161/08.
DR	N-PSDB; ABX76352.
XX	

PR Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer.
 XX
 PS Claim 27, Page 352, 453pp; English.

The invention relates to a method for detecting a lung cancer-associated transcript in a cell from a patient, comprising contacting a biological sample from the patient with a polynucleotide that selectively hybridizes to a sequence that is at least 80 % identical to a gene that exhibits increased or decreased expression in lung cancer samples. Lung cancer-associated polynucleotides and polypeptides are used for identifying a compound that modulates a lung cancer-associated polypeptide, for inhibiting proliferation of a lung cancer-associated cell to treat lung cancer in a patient and for treating a mammal having lung cancer by administering a modulatory compound identified. The methods are useful for treating lung cancer, such as small cell lung cancer, non-small cell lung cancer or other benign or precancerous lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides and polypeptides are useful for diagnostic purposes and as targets for screening for therapeutic compounds that modulate lung cancer, such as antibodies. Sequences ABUS6408-ABUS6745 represent lung cancer-associated polypeptides of the invention.

SQ Sequence 755 AA;

Query Match	100.0%	Score 3826	DB 6	Length 755
-------------	--------	------------	------	------------

Best Local Similarity 100.0%; Pred. No. 6.6e-300,
Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	LOEVHVSKEITGKISAA	RMW	CSA	AVD	IM	FL	LD	SN	VG	KS	FES	K	FA	IV	CG	LD	60																																	
Db	24	LOEVHVSKEITKISAA	RMW	CSA	AVD	IM	FL	LD	SN	VG	KS	FES	K	FA	IV	CG	LD	83																																	
Qy	61	SPEERVAGAFQES	TP	HL	EF	PL	DS	FT	Q	EV	K	IR	M	F	K	G	R	TE	LA	KY	L	H	R	GLP	120																										
Db	84	SPEERVAGAFQES	TP	HL	EF	PL	DS	FT	Q	EV	K	IR	M	F	K	G	R	TE	LA	KY	L	H	R	GLP	143																										
Qy	121	GGRNNSVPOILL	I	VT	D	G	K	S	Q	D	V	A	L	P	S	K	O	L	K	E	R	G	V	T	V	A	V	G	R	F	P	M	E	E	L	H	A	S	E	P	R	G	180								
Db	144	GGRNNSVPOILL	I	VT	D	G	K	S	Q	D	V	A	L	P	S	K	O	L	K	E	R	G	V	T	V	A	V	G	R	F	P	M	E	E	L	H	A	S	E	P	R	G	203								
Qy	181	HVLLAEQVEDATNG	L	F	S	T	L	S	S	A	I	C	S	A	T	P	D	C	R	V	E	H	P	C	E	H	R	L	E	M	V	R	E	P	A	G	N	A	P	C	M	R	G	240							
Db	204	HVLLAEQVEDATNG	L	F	S	T	L	S	S	A	I	C	S	A	T	P	D	C	R	V	E	H	P	C	E	H	R	L	E	M	V	R	E	P	A	G	N	A	P	C	M	R	G	263							
Qy	241	RRTLAVLAHACFP	Y	S	M	K	R	V	F	L	T	H	P	A	T	C	Y	T	T	C	P	G	P	D	S	O	P	C	O	N	G	T	C	V	E	B	G	L	D	G	Y	O	C	L	P	300					
Db	264	RRTLAVLAHACFP	Y	S	M	K	R	V	F	L	T	H	P	A	T	C	Y	T	T	C	P	G	P	D	S	O	P	C	O	N	G	T	C	V	E	B	G	L	D	G	Y	O	C	L	P	323					
Qy	301	IAFGSEANCALK	L	S	E	C	R	V	D	L	L	F	L	D	S	S	A	G	T	L	D	G	F	L	R	A	K	V	F	Y	K	R	V	R	A	V	L	S	D	S	R	A	R	V	360						
Db	324	IAFGSEANCALK	L	S	E	C	R	V	D	L	L	F	L	D	S	S	A	G	T	L	D	G	F	L	R	A	K	V	F	Y	K	R	V	R	A	V	L	S	D	S	R	A	R	V	383						
Qy	361	VATYSRELLVA	N	P	V	G	E	Y	O	V	P	D	L	W	S	D	I	P	R	G	E	T	L	L	G	S	A	L	R	O	A	E	E	G	S	A	T	R	T	G	420										
Db	384	VATYSRELLVA	N	P	V	G	E	Y	O	V	P	D	L	W	S	D	I	P	R	G	E	T	L	L	G	S	A	L	R	O	A	E	E	G	S	A	T	R	T	G	443										
Qy	421	RPRRVVLLT	S	E	S	E	D	E	V	G	P	A	R	H	A	R	E	L	L	L	G	V	S	E	A	V	R	A	E	L	E	I	T	G	S	P	K	V	M	Y	S	D	480								
Db	444	RPRRVVLLT	S	E	S	E	D	E	V	G	P	A	R	H	A	R	E	L	L	L	G	V	S	E	A	V	R	A	E	L	E	I	T	G	S	P	K	V	M	Y	S	D	503								
Qy	481	ODLFNOIP	E	L	O	G	K	L	C	S	R	O	P	G	C	R	T	O	A	L	D	V	F	M	L	T	S	A	S	V	G	E	N	F	A	O	M	O	S	F	V	R	S	C	A	L	O	F	E	V	540
Db	504	ODLFNOIP	E	L	O	G	K	L	C	S	R	O	P	G	C	R	T	O	A	L	D	V	F	M	L	T	S	A	S	V	G	E	N	F	A	O	M	O	S	F	V	R	S	C	A	L	O	F	E	V	563
Qy	541	NDVTVQVGLV	V	Y	G	S	Q																																												

QY 601 VQGGAPPGVPAVVVLTGGGAEADAAPQAQKLNNNGISVLTVGVPVLTSEGLRRLAGPRD 660
 DB 624 VQGGAPPGVPAVVVLTGGGAEADAAPQAQKLNNNGISVLTVGVPVLTSEGLRRLAGPRD 683
 QY 661 SLIHVAAYADLRHYHGVLTLEMLCGEAKOPVNLCKPSPCMNBSGCVLQNSYRCKCRDGM 720
 DB 684 SLIHVAAYADLRHYHGVLTLEMLCGEAKOPVNLCKPSPCMNBSGCVLQNSYRCKCRDGM 743
 QY 721 GPHCENRFLRRP 732
 DB 744 GPHCENRFLRRP 755

RESULT 4
 ADB80523
 ID ADB80523 standard; protein; 755 AA.
 XX ADB80523;
 AC ADB80523;
 DT 04-DEC-2003 (first entry)
 XX
 DE Ovarian cancer-associated protein #44.
 XX
 KM cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
 KM post-operative chemotherapy; radiation therapy; tumour prognosis;
 KM pre-cancerous lesion detection.
 XX
 OS Homo sapiens.
 XX
 PN MO2002102235-AA.
 XX
 PD 27-DEC-2002.
 XX
 PF 18-JUN-2002; 2002MO-US019297.
 XX
 PR 18-JUN-2001; 2001US-0299234P.
 PR 27-APR-2001; 2001US-0315287P.
 PR 05-SEP-2001; 2001US-0317544P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 12-APR-2002; 2002US-0372246P.
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Mack DH, Glash KC;
 XX
 DR MPI; 2003-167431/16.
 XX
 PT N-PSDB; ADB80522.
 XX
 PT Detecting an ovarian cancer-associated transcript in a cell from a
 PT patient, comprises contacting a biological sample from the patient with a
 PT polynucleotide that hybridizes to an ovarian cancer gene.
 XX
 PS Claim 13; Page 304-305; 332pp; English.
 XX
 CC The invention relates to a method of detecting an ovarian cancer-
 CC associated transcript in a cell from a patient, by contacting a
 CC biological sample from the patient with a polynucleotide that selectively
 CC hybridizes to a sequence at least 80% identical to any of one of 80
 CC nucleic acid sequences given in the specification. The method is useful
 CC in diagnosing ovarian cancer and in identifying and using agents and/or
 CC targets that inhibit ovarian cancer. The nucleic acid molecule,
 CC polypeptide and the antibody may also be used in detecting ovarian
 CC cancers, monitoring and early detection of relapse following treatment,
 CC monitoring response to therapy, selecting patients for post-operative
 CC chemotherapy or radiation therapy, in selecting mode of therapy,
 CC determining tumour prognosis, early detection of pre-cancerous lesions,
 CC and as vaccines. This sequence corresponds to one of the proteins used
 CC for the detection method of the invention.
 CC
 XX
 SQ Sequence 755 AA;
 Query Match 100.0%; Score 3826; DB 7; Length 755;
 Best Local Similarity 100.0%; Pred. No. 6.6e-300;

Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LQEVHSKETTIGISAASQMWCSAAVDIMFLDSSNSVKGKSPERSKFAITVCGLDI 60
 DB 24 LQEVHSKETTIGISAASQMWCSAAVDIMFLDSSNSVKGKSPERSKFAITVCGLDI 83
 QY 61 SPEERVGAFOFSSPHLEFPIDSFSTQOEVARIKRMVFKGRTETELAKYLLHRGLP 120
 DB 84 SPEERVGAFOFSSPHLEFPIDSFSTQOEVARIKRMVFKGRTETELAKYLLHRGLP 143
 QY 121 GGRNASVPQILITVTDGKSQGDVALPSKOLKRGVTVPVGVFRPMEELHLLASPRQ 180
 DB 144 GGRNASVPQILITVTDGKSQGDVALPSKOLKRGVTVPVGVFRPMEELHLLASPRQ 203
 QY 181 HYLAEQVEDATNGLPSTLSSAICSSATPDGRVAHPCEHRTLELVREFPAGNAPCMRGS 240
 DB 204 HYLAEQVEDATNGLPSTLSSAICSSATPDGRVAHPCEHRTLELVREFPAGNAPCMRGS 263
 QY 241 RRTTLAVLAHCPFYGMKRVFLTHPATCYRTTCPGFCDSPQCONGGTCVBEGLDGYOCLCP 300
 DB 264 RRTTLAVLAHCPFYGMKRVFLTHPATCYRTTCPGFCDSPQCONGGTCVBEGLDGYOCLCP 323
 QY 301 IAFGGEANCALKLSIECRVDLFLDSSAGTLDGFLRAKVVKRFPVAVLSEDSRVRG 360
 DB 324 IAFGGEANCALKLSIECRVDLFLDSSAGTLDGFLRAKVVKRFPVAVLSEDSRVRG 383
 QY 361 VATYSRELLVAVPVEGYODPDLVWSLDGIPPRGGPTLTGSLARQAAERFGSATRTGQD 420
 DB 384 VATYSRELLVAVPVEGYODPDLVWSLDGIPPRGGPTLTGSLARQAAERFGSATRTGQD 443
 QY 421 RPRRVVVLITESHSEDEVAGPARHARARELLLVGSEAVNABELEITGSPGHVWYSDP 480
 DB 444 RPRRVVVLITESHSEDEVAGPARHARARELLLVGSEAVNABELEITGSPGHVWYSDP 503
 QY 481 QDLFNOIPELAGKLSRORPGCRTOALDVFELDTASAVGPNFPMOMOSFVRSCLAQEPV 540
 DB 504 QDLFNOIPELAGKLSRORPGCRTOALDVFELDTASAVGPNFPMOMOSFVRSCLAQEPV 563
 QY 541 NPDVTQVGLVYVGSQVQTAFGIDTKPTRAAMLRAISQAPVYLGVSAGTALHITDKWMT 600
 DB 564 NPDVTQVGLVYVGSQVQTAFGIDTKPTRAAMLRAISQAPVYLGVSAGTALHITDKWMT 623
 QY 601 VQGGAPPGVPAVVVLTGGGAEADAAPQAQKLNNNGISVLTVGVPVLTSEGLRRLAGPRD 660
 DB 624 VQGGAPPGVPAVVVLTGGGAEADAAPQAQKLNNNGISVLTVGVPVLTSEGLRRLAGPRD 683
 QY 661 SLIHVAAYADLRHYHGVLTLEMLCGEAKOPVNLCKPSPCMNBSGCVLQNSYRCKCRDGM 720
 DB 684 SLIHVAAYADLRHYHGVLTLEMLCGEAKOPVNLCKPSPCMNBSGCVLQNSYRCKCRDGM 743
 QY 721 GPHCENRFLRRP 732
 DB 744 GPHCENRFLRRP 755

RESULT 5
 ADN38806
 ID ADN38806 standard; protein; 755 AA.
 XX ADN38806;
 AC ADN38806;
 DT 17-JUN-2004 (first entry)
 XX
 DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:124.
 XX
 KM Human; differential expression; cancer; angiogenic disorder;
 KM fibrotic disorder; porriasis; ischaemia; heart disease; atherosclerosis;
 KM inflammatory disease; autoimmune disease;
 KM retinal neovascularisation syndrome; scarring; uterine fibroid;
 KM detection; diagnosis; prognosis; drug screening; drug targeting;
 KM wound healing; contraception; cytostatic; cardiac; immunomodulatory;
 KM vulnerary; gene therapy; vaccine.
 XX

OS Homo sapiens.
 XX WO2003042661-A2.
 XX
 XX 22-MAY-2003.
 XX
 PF 13-NOV-2002; 2002WO-US036810.
 XX
 PR 13-NOV-2001; 2001US-0350666P.
 PR 21-NOV-2001; 2001US-0332464P.
 PR 29-NOV-2001; 2001US-0334393P.
 PR 03-DEC-2001; 2001US-0335394P.
 PR 14-DEC-2001; 2001US-0340376P.
 PR 08-JAN-2002; 2002US-0347211P.
 PR 10-JAN-2002; 2002US-0347349P.
 PR 08-FEB-2002; 2002US-0355250P.
 PR 13-FEB-2002; 2002US-0356714P.
 PR 20-FEB-2002; 2002US-0359077P.
 PR 29-MAR-2002; 2002US-0368809P.
 PR 04-APR-2002; 2002US-0370110P.
 PR 12-APR-2002; 2002US-0372246P.
 PR 05-JUN-2002; 2002US-038614P.
 PR 16-JUL-2002; 2002US-0396839P.
 PR 22-JUL-2002; 2002US-0397755P.
 PR 22-JUL-2002; 2002US-0397845P.
 PR 09-SEP-2002; 2002US-0409450P.
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Afar D, Aziz N, Gineburg WM, Gish KC, Glynn R, Hevezi PA;
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlocnik A;
 XX
 DR WPI; 2003-468649/44.
 XX N-PSDB; ADN38805.
 XX
 PT Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.
 XX
 PS Claim 12; SEQ ID NO 124; 1385bp; English.
 XX
 CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularization syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a polypeptide of the invention.
 XX
 XX Sequence 755 AA;
 SO

Query Match 100.0%; Score 3826; DB 7; Length 755;
 Best Local Similarity 100.0%; Pred. No. 6,6e-300; Indels 0; Gaps 0;
 Matches 732; Conservative 0; Mismatches 0;

QY 1 LOEVHYSKETIGTISAAKMMCSAAVDIMFLDGSNSVGSGPERSKHPAIVTCGLDI 60
 XX
 DB 24 LOEVHYSKETIGTISAAKMMCSAAVDIMFLDGSNSVGSGPERSKHPAIVTCGLDI 83
 XX
 QY 61 SPERVAVGAFQFSSTPHLEPPLDSFSTQGEVKARIKRVKGRTEETELAKTLHRLGP 120
 XX
 DB 84 SPERVAVGAFQFSSTPHLEPPLDSFSTQGEVKARIKRVKGRTEETELAKTLHRLGP 143
 XX

QY 121 GGRNASVPOIILIIYTDGKSQGDVALPSKQLKRGVTVFVAVRPFPRMELHMLASEPRQ 180
 |||||
 DB 144 GGRNASVPOIILIIYTDGKSQGDVALPSKQLKRGVTVFVAVRPFPRMELHMLASEPRQ 203
 |||||
 QY 181 HTLLABQVEDATNGLPSTLSSAICSSATPDCCREAHPCENHRTLEKMBEPAQNAQWRGS 240
 |||||
 DB 204 HTLLABQVEDATNGLPSTLSSAICSSATPDCCREAHPCENHRTLEKMBEPAQNAQWRGS 263
 |||||
 QY 241 RRTLLVLAHCPFPYMKRVFLTHPATCYRTTCRPGCDSPQONGGCVPEGLDGYQCLCP 300
 |||||
 DB 264 RRTLLVLAHCPFPYMKRVFLTHPATCYRTTCRPGCDSPQONGGCVPEGLDGYQCLCP 323
 |||||
 QY 301 IAFGSEANCAIKLISECRVDLFLDSSAGTTLDFLPAKVFKFVPAVLSSEDSRARVG 360
 |||||
 DB 324 IAFGSEANCAIKLISECRVDLFLDSSAGTTLDFLPAKVFKFVPAVLSSEDSRARVG 383
 |||||
 QY 361 VATTYRELLVAVPVEEYQDPDLVMSLDGIPFRGGPTLTGSLARQAAERGFSSATRTGQD 420
 |||||
 DB 384 VATTYRELLVAVPVEEYQDPDLVMSLDGIPFRGGPTLTGSLARQAAERGFSSATRTGQD 443
 |||||
 QY 421 RRRRVVVLITESHSEDEVAGPARHARRELLLLGVSEAVRALEBEITGSPGHVWVYSDP 480
 |||||
 DB 444 RRRRVVVLITESHSEDEVAGPARHARRELLLLGVSEAVRALEBEITGSPGHVWVYSDP 503
 |||||
 QY 481 QDLFNQIPELQKLCSPRPGCRTOALDIVFMLDTSASVGPENFAQOMQSFVNSCALQFEV 540
 |||||
 DB 504 QDLFNQIPELQKLCSPRPGCRTOALDIVFMLDTSASVGPENFAQOMQSFVNSCALQFEV 563
 |||||
 QY 541 NPDVTQVGLVYVGSQVQTAFFGLDTPKTPRAMLRAISQADYLGCVSAGTALLHIYDKVMT 600
 |||||
 DB 564 NPDVTQVGLVYVGSQVQTAFFGLDTPKTPRAMLRAISQADYLGCVSAGTALLHIYDKVMT 623
 |||||
 QY 601 VORGARPGVPAKVVVLTGCRGAEDAAVPAQKLRNNGISVLYVGVGPVYSEGLRRLAGPRD 660
 |||||
 DB 624 VORGARPGVPAKVVVLTGCRGAEDAAVPAQKLRNNGISVLYVGVGPVYSEGLRRLAGPRD 683
 |||||
 QY 661 SLIHVAAYADLHYHDVLEIEMLCGAKQPVNLCKPSPCMNBSGCVLQNSYRCKRGDME 720
 |||||
 DB 684 SLIHVAAYADLHYHDVLEIEMLCGAKQPVNLCKPSPCMNBSGCVLQNSYRCKRGDME 743
 |||||
 QY 721 GPHCENRFLRRP 732
 |||||
 DB 744 GPHCENRFLRRP 755
 |||||

RESULT 6
 ADL70227
 ID ADL70227 standard; protein; 755 AA.
 XX
 AC ADL70227;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE Colon cancer marker Colo2p2 full-length protein.
 XX
 KM Colo2p2: colon cancer; marker; diagnosis; cytostatic; vaccine;
 KM gene therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..23
 FT Protein /label= Signal peptide
 FT /label= Secreted protein
 FT /note= "Also referred to in Claim 10"
 XX
 PN MO2004018648-A2.
 XX
 PD 04-MAR-2004.
 XX
 PF 26-AUG-2003; 2003WO-US027086.
 XX

PR 26-AUG-2002; 2002US-00229345.
 PR 27-AUG-2002; 2002US-0406296P.
 PR 18-OCT-2002; 2002US-00274177.
 XX
 PA (UYCA-) UNIV CASE WESTERN RESERVE.
 XX
 PI Markowitz SD;
 XX
 DR WPI; 2004-226834/21.
 XX
 PT Inhibiting the growth or proliferation of a colon neoplasia in a subject,
 PT for treating colon cancer, comprises administering to the subject an
 PT agent that decreases the amount of a polypeptide present in or produced
 PT by the colon neoplasia.
 PS
 PS Claim 10; SEQ ID NO 14; 125pp; English.
 XX
 CC The present sequence is the protein sequence of full-length ColoUp2, a
 CC molecular marker of colon neoplasia. ColoUp2 was identified by expression
 CC microarray profiling of genes showing differential expression in normal
 CC colon and metastatic colon cancer. ColoUp2 transcript is essentially
 CC undetectable in normal colon epithelial strips, in normal and in colonic
 CC muscle, but expression is clearly detectable in premalignant colon
 CC adenomas and in 90% of Dukes stage B (early negative colon cancers),
 CC Dukes stage C (node positive colon cancer), Dukes stage D (primary colon
 CC cancers with associated metastatic spread) and in colon cancer liver
 CC metastasis. Expression is also detectable in colon cancer cell lines and
 CC in colon cancer xenografts grown in athymic mice. ColoUp2 is secreted
 CC from both the basolateral and apical surfaces of intestinal cells, and
 CC can be found in the blood in 2 different forms, a full-length secreted
 CC form and a C-terminal fragment. The molecular marker is useful in the
 CC detection or diagnosis of colon neoplasia and for categorizing the
 CC neoplastic state of a patient. Methods for inhibiting the growth or
 CC proliferation of a colon neoplasia involve the use of short interfering
 CC RNA (siRNA) or antisense oligonucleotides, or an agent that binds to and
 CC antagonizes the polypeptide. Molecular markers can also be used to target
 CC therapeutic agents to cells of a colon neoplasia, and to screen and
 CC identify candidate therapeutic agents.
 XX
 XX Sequence 755 AA;
 SQ
 Query Match 100.0%; Score 3826; DB 8; Length 755;
 Best Local Similarity 100.0%; Pred. No. 6, 6e-300;
 Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 421 RRRRVVTLTSHSEDEVAGPARHARRELLILGVGSEAVRAELEBITGSPRHVMVYSDP 480
 DB 444 RRRRVVTLTSHSEDEVAGPARHARRELLILGVGSEAVRAELEBITGSPKHVMVYSDP 503
 QY 481 QDLFNOIPELQKLSRORPGCRTOALDVFMLDTSASVGPENFOMOSFVNSCALQFEV 540
 DB 504 QDLFNOIPELQKLSRORPGCRTOALDVFMLDTSASVGPENFOMOSFVNSCALQFEV 563
 QY 541 NPDVTQVGLVYVGSQVOTAFGLDTKPTRAAMLRAISOAPYLGVGSAGTALHIYDKVMT 600
 DB 564 NPDVTQVGLVYVGSQVOTAFGLDTKPTRAAMLRAISOAPYLGVGSAGTALHIYDKVMT 623
 QY 601 VORGARPGVPKAVVVLTGGRGADAAVPAQKLRNNGISVLYVGVGPVUSGILRLAGPRD 660
 DB 624 VORGARPGVPKAVVVLTGGRGADAAVPAQKLRNNGISVLYVGVGPVUSGILRLAGPRD 683
 QY 661 SLIHVAAYADLRHYDDVLIEMLCGRAKOPVNLCKRSPCKNBSGCVLQNGSYRCKCRDGE 720
 DB 684 SLIHVAAYADLRHYDDVLIEMLCGRAKOPVNLCKRSPCKNBSGCVLQNGSYRCKCRDGE 743
 QY 721 GPHCENRFLRRP 732
 DB 744 GPHCENRFLRRP 755
 RESULT 7
 ADNS9619
 ID ADNS9619 strand; protein; 755 AA.
 XX
 AC ADNS9619;
 XX
 DT 01-JUL-2004 (first entry)
 XX
 DB Colon neoplasia full-length ColoUp2 protein.
 XX
 KW molecular marker; neoplastic state; colon neoplasia; ColoUp1; ColoUp2;
 KW colon cancer.
 OS
 OS Unidentified.
 XX
 PN MO2004018647-A2.
 XX
 PD 04-MAR-2004.
 XX
 PF 26-AUG-2003; 2003WO-US027085.
 XX
 PR 26-AUG-2002; 2002US-00229345.
 PR 18-OCT-2002; 2002US-00274177.
 XX
 PA (UYCA-) UNIV CASE WESTERN RESERVE.
 PI Markowitz SD;
 XX
 DR WPI; 2004-419536/39.
 XX
 PT Detecting likelihood of subject to have colon neoplasia, involves
 PT obtaining biological sample from subject, detecting one or more of
 PT ColoUp1 and ColoUp2 polypeptides, presence of polypeptide indicating
 PT colon neoplasia.
 PS
 PS Claim 11; SEQ ID NO 14; 118pp; English.
 XX
 CC The invention relates to novel molecular markers for categorizing the
 CC neoplastic state of a patient. The invention further comprises a method
 CC for detecting whether a subject is likely to have a colon neoplasia. The
 CC method comprises: obtaining a biological sample from the subject;
 CC detecting one or more polypeptides selected from among: one or more
 CC ColoUp1 polypeptides and one or more ColoUp2 polypeptides, wherein the
 CC presence of said one or more polypeptides is indicative of colon
 CC neoplasia. The invention further comprises a kit for carrying out the
 CC said method. The methods and compositions are useful for detecting colon
 CC cancer. This sequence represents a full-length ColoUp2 protein of the
 CC invention.

XX Sequence 755 AA;
 SQ Query Match 100.0%; Score 3826; DB 8; Length 755;
 Best Local Similarity 100.0%; Pred. No. 6,6e-300;
 Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 LOEVHVSKEITGKISASMMWCSAAVDIMPLDGSNSVGKGSFERSKFAITVCDGLDI 60
 24 LOEVHVSKEITGKISASMMWCSAAVDIMPLDGSNSVGKGSFERSKFAITVCDGLDI 83

QY 61 SPERVAVGAFOFSSSTPHLEFPDLSFSTQOEYKARIMRVFKGRTETELAKYLLHRGUP 120
 DB 84 SPERVAVGAFOFSSSTPHLEFPDLSFSTQOEYKARIMRVFKGRTETELAKYLLHRGUP 143

QY 121 GGRNASVPOILLIIVTDKSGQDVALLPSKQIKERGVTVFVAVGRFPFMEELHALASPRQ 180
 DB 144 GGRNASVPOILLIIVTDKSGQDVALLPSKQIKERGVTVFVAVGRFPFMEELHALASPRQ 203

QY 181 HYLAAQVDAATNGLFSTLSSSAICSSATPDCVKAHPCBHRTLEWVRFAGNAPCMRGS 240
 DB 204 HYLAAQVDAATNGLFSTLSSSAICSSATPDCVKAHPCBHRTLEWVRFAGNAPCMRGS 263

QY 241 RRTLAVALAHCPEYSKRWVFLTHPATCYRTTCPCDSQPCQNGGTCVEBGLDGYCQLCP 300
 DB 264 RRTLAVALAHCPEYSKRWVFLTHPATCYRTTCPCDSQPCQNGGTCVEBGLDGYCQLCP 323

QY 301 LAFGSEANCAIKLISLECRVDLLFLDSSAGTTIDGFLRAKRVYKRFVRAVLSDSRARV 360
 DB 324 LAFGSEANCAIKLISLECRVDLLFLDSSAGTTIDGFLRAKRVYKRFVRAVLSDSRARV 383

QY 361 VATYSBELVAVAVGVGYQVDPDLVMSLDGIPRPGGPTLGLSALROAEEFGSATGTGD 420
 DB 384 VATYSBELVAVAVGVGYQVDPDLVMSLDGIPRPGGPTLGLSALROAEEFGSATGTGD 443

QY 421 RRRRVVLLTESHSEDEVAAPAHARARELLILGVSEAVRALEBEITGSPKHWYSDP 480
 DB 444 RRRRVVLLTESHSEDEVAAPAHARARELLILGVSEAVRALEBEITGSPKHWYSDP 503

QY 481 QDLFNOIPELOGKLCGRORPGCCTOALDIVFMLDTSASVGENFAQMGSPVNSCALQFV 540
 DB 504 QDLFNOIPELOGKLCGRORPGCCTOALDIVFMLDTSASVGENFAQMGSPVNSCALQFV 563

QY 541 NPDVNOGVGVVSSQVOTAFGLDTKPTRAMLAISQAPVLGVSAGATLLHIYKWT 600
 DB 564 NPDVNOGVGVVSSQVOTAFGLDTKPTRAMLAISQAPVLGVSAGATLLHIYKWT 623

QY 601 VORGAAPVPAVVLTGGRGADAAVPAQKLANNGISVLYVGVGVLSBGLRRLAGPRD 660
 DB 624 VORGAAPVPAVVLTGGRGADAAVPAQKLANNGISVLYVGVGVLSBGLRRLAGPRD 683

QY 661 SLIHVAAYADLRHYQDVLIEWLGEAKOPVNLCKPSPCMNBSGCVLQNSYRCKRDGE 720
 DB 684 SLIHVAAYADLRHYQDVLIEWLGEAKOPVNLCKPSPCMNBSGCVLQNSYRCKRDGE 743

QY 721 GPHCENRPLRRP 732
 DB 744 GPHCENRPLRRP 755

RESULT 8
 ADN39868
 ID ADN39868 standard; protein; 755 AA.
 XX
 AC ADN39868;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE Cancer/angiogenesis/fibrosis-related polypeptide. SEQ ID NO: C238.
 XX
 KW Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;

KW retinal neovascularization syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiac; immunomodulatory;
 KW vulnery; gene therapy; vaccine.
 XX Homo sapiens.
 XX MO2003042661-A2.
 XX PD 22-MAY-2003.
 XX
 XX 13-NOV-2002; 2002WO-US036810.
 XX PF
 XX 13-NOV-2001; 2001US-0350666P.
 XX PR 21-NOV-2001; 2001US-0332464P.
 XX PR 29-NOV-2001; 2001US-0334393P.
 XX PR 03-DEC-2001; 2001US-0335394P.
 XX PR 14-DEC-2001; 2001US-0340376P.
 XX PR 08-JAN-2002; 2002US-0347211P.
 XX PR 10-JAN-2002; 2002US-0347349P.
 XX PR 08-FEB-2002; 2002US-0355250P.
 XX PR 13-FEB-2002; 2002US-0356714P.
 XX PR 20-FEB-2002; 2002US-0359777P.
 XX PR 29-MAR-2002; 2002US-036809P.
 XX PR 04-APR-2002; 2002US-0370110P.
 XX PR 12-APR-2002; 2002US-0372246P.
 XX PR 05-JUN-2002; 2002US-0386614P.
 XX PR 16-JUL-2002; 2002US-0396839P.
 XX PR 22-JUL-2002; 2002US-0397757P.
 XX PR 09-SEP-2002; 2002US-0409450P.
 XX
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 XX PA
 XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevez PA;
 XX PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
 XX
 XX WPI: 2003-468649/44.
 XX DR N-PSDB; ADN39651.
 XX
 XX Determining the presence or absence of a pathological cell in a patient,
 XX PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 XX PT a nucleic acid in a biological sample.
 XX
 XX
 XX Claim 12; SEQ ID NO C238; 1385pp; English.
 XX PS
 XX CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 XX CC whose expression is upregulated or downregulated in specific cancers or
 XX CC other diseases such as angiogenic or fibrotic disorders, and to methods
 XX CC of determining the presence or absence of a pathological cell in a
 XX CC patient by detecting a nucleic acid at least 80% identical to those of
 XX CC the invention or by detecting a polypeptide of the invention. The
 XX CC invention also relates to expression vectors and host cells comprising a
 XX CC nucleic acid of the invention; antibodies which specifically bind a
 XX CC polypeptide of the invention; use of such antibodies for drug targeting;
 XX CC and methods of screening for modulators of activity or expression of the
 XX CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 XX CC antibodies and methods are useful for diagnosing, prognosing and treating
 XX CC cancer and other conditions such as psoriasis, ischemia, heart disease,
 XX CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 XX CC neovascularization syndromes, scarring and uterine fibroids. They may
 XX CC also be useful in wound healing and in contraception. The present
 XX CC sequence represents a polypeptide of the invention.
 XX
 XX
 XX SQ Sequence 755 AA;
 XX
 XX Query Match 99.9%; Score 3822; DB 7; Length 755;
 XX Best Local Similarity 99.9%; Pred. No. 1.4e-299;
 XX Matches 731; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LOEVHVSKEITGKISASMMWCSAAVDIMPLDGSNSVGKGSFERSKFAITVCDGLDI 60
 DB 24 LOEVHVSKEITGKISASMMWCSAAVDIMPLDGSNSVGKGSFERSKFAITVCDGLDI 83

```

QY 61 SPERRVGAFOFSSSTPHLEPPLDSFSTQGEVKARIKRMVKGGRTEITELAKYLHRLGP 120
Db 84 SPERRVGAFOFSSSTPHLEPPLDSFSTQGEVKARIKRMVKGGRTEITELAKYLHRLGP 143
QY 121 GGRNASVPQILLITVDGKSQGDVALPSKOLKRGVTVAVGVPFRMEIHALASEPRQ 180
Db 144 GGRNASVPQILLITVDGKSQGDVALPSKOLKRGVTVAVGVPFRMEIHALASEPRQ 203
QY 181 HVLAEQVEDATNGLFSTLSSSAICSSATPDCRVEAHPCHRTLEMYREPAAGNAPCMRGS 240
Db 204 HVLAEQVEDATNGLFSTLSSSAICSSATPDCRVEAHPCHRTLEMYREPAAGNAPCMRGS 263
QY 241 RRTLAIVLAHCPFYSWKRVFLTHPATCYRTTCGPGCDSPQONGTCTVPEGIDGYQCLCP 300
Db 264 RRTLAIVLAHCPFYSWKRVFLTHPATCYRTTCGPGCDSPQONGTCTVPEGIDGYQCLCP 323
QY 301 LAFGEANCAKLSLECRVDLLFLDSSAGTTLDFLRAKVFYKRVRAVLSEDSRARVG 360
Db 324 LAFGEANCAKLSLECRVDLLFLDSSAGTTLDFLRAKVFYKRVRAVLSEDSRARVG 383
QY 361 VATYSRELLVAVPGEYQDVLDVMSLDGIPFRGSPITLTSALRQAERGFSGATRTGQD 420
Db 384 VATYSRELLVAVPGEYQDVLDVMSLDGIPFRGSPITLTSALRQAERGFSGATRTGQD 443
QY 421 RPRRVVLLTSHSEDEBAGAPAHARARELLLLGVGSEAVRAELEBITGSPKRWVYSDP 480
Db 444 RPRRVVLLTSHSEDEBAGAPAHARARELLLLGVGSEAVRAELEBITGSPKRWVYSDP 503
QY 481 QDLFNOIPELOGLCSRRPGCTONLDVFMLDTSASVGPENFAMQSVNRCALOFVY 540
Db 504 QDLFNOIPELOGLCSRRPGCTONLDVFMLDTSASVGPENFAMQSVNRCALOFVY 563
QY 541 NPDTVOGLVYVSQVOTAFGLDTKPTRAAMLRAISQAPYLGVSAGTALLHIYDKWT 600
Db 564 NPDTVOGLVYVSQVOTAFGLDTKPTRAAMLRAISQAPYLGVSAGTALLHIYDKWT 623
QY 601 VORGAPGVPAKAVVLTGGRGAEDAAVPAQKLRNNGISVLVVGVPVLSBGLRRLAGPRD 660
Db 624 VORGAPGVPAKAVVLTGGRGAEDAAVPAQKLRNNGISVLVVGVPVLSBGLRRLAGPRD 683
QY 661 SLIHVAAYDLRHOVLTEMLGGEAKOPNLCBSPCMWEGSCVQNSYRCKCKDGM 720
Db 684 SLIHVAAYDLRHOVLTEMLGGEAKOPNLCBSPCMWEGSCVQNSYRCKCKDGM 743
QY 721 GPHCENRFLRRP 732
Db 744 GPHCENRFLRRP 755

```

```

RESULT 9
ABG61891
ID ABG61891 standard; protein; 807 AA.
XX
AC ABG61891;
XX
DT 15-AUG-2002 (first entry)
XX
DE Prostate cancer-associated protein #92.
XX
KM Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.
XX
OS Mammalia.
XX
PN WO200230268-A2.
XX
PD 18-APR-2002.
XX
PF 12-OCT-2001; 2001WO-US032045.
XX
PR 13-OCT-2000; 2000US-00687576.
PR 08-DEC-2000; 2000US-00733288.
PR 08-DEC-2000; 2000US-00733742.

```

```

PR 24-JAN-2001; 2001US-0263957P.
PR 16-MAR-2001; 2001US-0276791P.
PR 16-MAR-2001; 2001US-0276888P.
PR 06-APR-2001; 2001US-0281922P.
PR 24-APR-2001; 2001US-0286214P.
PR 30-APR-2001; 2001US-00847046.
PR 04-MAY-2001; 2001US-0288589P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Gish KC, Mack DH, Wilson KE, Afar D, Heyezi P;
XX
DR WPI; 2002-471335/50.
XX
DR N-PSDB; ABK92207.
XX
PT Detecting a prostate cancer-associated transcript in a cell in a patient,
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
PT by determining if prostate cancer-associated genes are expressed in a
PT prostate tissue.
XX
PS Claim 27; Page 377; 436pp; English.
XX
CC The present invention relates to methods of detecting a prostate cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with prostate cancer-
CC associated polynucleotides (designated PC genes) that selectively
CC hybridize to a sequence that is at least 80% identical to them. The
CC prostate cancer-associated polynucleotide sequences are differentially
CC expressed in prostate tumour tissue or in prostate cancer and are derived
CC from the tissues of various organisms such as humans or other mammals
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
CC associated genes are useful for diagnosing or treating prostate cancer,
CC as well as for identifying modulators of prostate cancer or agents that
CC inhibit prostate cancer. The nucleic acid sequences are particularly
CC useful in gene therapy, as a vaccine or in antisense applications.
CC ABG61800-ABG61944 represent prostate cancer-associated proteins
XX
SQ Sequence 807 AA;

```

```

Query Match 99.3%; Score 3799; DB 5; Length 807;
Best Local Similarity 100.0%; Pred. No. 1,1e-297;
Matches 727; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LOEVAVSKETIGKISAASKMMCSAAVDIMFLDGSNSVGKSPERSKHFAITVCDGLDI 60
Db 24 LOEVAVSKETIGKISAASKMMCSAAVDIMFLDGSNSVGKSPERSKHFAITVCDGLDI 83
QY 61 SPERRVGAFOFSSSTPHLEPPLDSFSTQGEVKARIKRMVKGGRTEITELAKYLHRLGP 120
Db 84 SPERRVGAFOFSSSTPHLEPPLDSFSTQGEVKARIKRMVKGGRTEITELAKYLHRLGP 143
QY 121 GGRNASVPQILLITVDGKSQGDVALPSKOLKRGVTVAVGVPFRMEIHALASEPRQ 180
Db 144 GGRNASVPQILLITVDGKSQGDVALPSKOLKRGVTVAVGVPFRMEIHALASEPRQ 203
QY 181 HVLAEQVEDATNGLFSTLSSSAICSSATPDCRVEAHPCHRTLEMYREPAAGNAPCMRGS 240
Db 204 HVLAEQVEDATNGLFSTLSSSAICSSATPDCRVEAHPCHRTLEMYREPAAGNAPCMRGS 263
QY 241 RRTLAIVLAHCPFYSWKRVFLTHPATCYRTTCGPGCDSPQONGTCTVPEGIDGYQCLCP 300
Db 264 RRTLAIVLAHCPFYSWKRVFLTHPATCYRTTCGPGCDSPQONGTCTVPEGIDGYQCLCP 323
QY 301 LAFGEANCAKLSLECRVDLLFLDSSAGTTLDFLRAKVFYKRVRAVLSEDSRARVG 360
Db 324 LAFGEANCAKLSLECRVDLLFLDSSAGTTLDFLRAKVFYKRVRAVLSEDSRARVG 383
QY 361 VATYSRELLVAVPGEYQDVLDVMSLDGIPFRGSPITLTSALRQAERGFSGATRTGQD 420
Db 384 VATYSRELLVAVPGEYQDVLDVMSLDGIPFRGSPITLTSALRQAERGFSGATRTGQD 443
QY 421 RPRRVVLLTSHSEDEBAGAPAHARARELLLLGVGSEAVRAELEBITGSPKRWVYSDP 480

```

DB 444 RRRRVVLLTSHSEDEVAGPARHARARELLLLGVSEAVRALEBEITGSPKHWVWYSDP 503
 QY 481 QDLFNOIPELOGLKCSRRPGCRTOALDVFMLDTSASVGPENFAOMOSFVMSCALQFEV 540
 DB 504 QDLFNOIPELOGLKCSRRPGCRTOALDVFMLDTSASVGPENFAOMOSFVMSCALQFEV 563
 QY 541 NPDVTOVGLVVGSGVQVTFAGLDTKPTRAMLRATISQAPYLGVSAGTALLHIYDKVMT 600
 DB 564 NPDVTOVGLVVGSGVQVTFAGLDTKPTRAMLRATISQAPYLGVSAGTALLHIYDKVMT 623
 QY 601 VORGARPGVPKAVVVLTTGGRGAEDAAVPAQKLRNNGISVLLVVGVPVLSBGLRRLAGPRD 660
 DB 624 VORGARPGVPKAVVVLTTGGRGAEDAAVPAQKLRNNGISVLLVVGVPVLSBGLRRLAGPRD 683
 QY 661 SLIHVAAYVADLRHYQDVLIEMLCGEAKOPVNLCKPSPCMNEGSCVLTQNGSYRCKCRDGE 720
 DB 684 SLIHVAAYVADLRHYQDVLIEMLCGEAKOPVNLCKPSPCMNEGSCVLTQNGSYRCKCRDGE 743
 QY 721 GPHCENR 727
 DB 744 GPHCENR 750

RESULT 10

ABP54687
 ID ABP54687 standard; protein; 807 AA.

AC ABP54687;

DT 30-DEC-2002 (first entry)

DE Metastatic colorectal cancer-associated polypeptide.

KM Colorectal cancer; metastasis; differential expression; cytostatic;
 diagnosis; gene therapy; vaccine.

OS Homo sapiens.

PN W0200268677-A2.

PD 06-SRP-2002.

PF 27-FEB-2002; 2002MO-US006001.

PR 27-FEB-2001; 2001US-0272206P.

PR 02-APR-2001; 2001US-0281149P.

PR 17-APR-2001; 2001US-0284555P.

PA (EOSB-) EOS BIOTECHNOLOGY INC.

PA (UYCA-) UNIV CASE WESTERN RESERVE.

PI Mack DH, Markowitz SD;

PI MPI; 2002-698677/75.

DR N-PSDB; AB081552.

XX New genes that are up- or down-regulated in colorectal cancer, useful for

XX diagnosing colorectal cancer in a subject, or for identifying modulators

XX of colorectal cancer-associated proteins and genes for treating

XX colorectal cancer.

XX Claim 8; Page 249; 260pp; English.

XX The present sequence is the protein sequence of a human polypeptide

XX encoded by a gene that exhibits increased expression in colon cancer-

XX derived metastases compared to normal colon tissue. It is an example of

XX claimed polypeptides that are encoded by genes which are differentially

XX expressed in metastatic colorectal cancer cells. Such polypeptides are

XX useful in diagnostic and prognostic assays, for raising antibodies useful

XX e.g. in immunotherapy, and in screening for modulator compounds of

XX therapeutic value

SQ Sequence 807 AA;

Query Match 99.3%; Score 3799; DB 5; Length 807;

Best Local Similarity 100.0%; Pred. No. 1.1e-297; Mismatches 0; Gaps 0;

Matches 727; Conservative 0; Indels 0; Gaps 0;

QY 1 LOEVHSKETIKISASAKMMWCSAAVDMIFLLDGSNSVGKGSFERSKHFAITVCDGLDI 60
 DB 24 LOEVHSKETIKISASAKMMWCSAAVDMIFLLDGSNSVGKGSFERSKHFAITVCDGLDI 83
 QY 61 SPEERVGAFOFSSTPHLEFPIDSFSTQOEVARIKRMVFKGRTETELALXYLHRLGP 120
 DB 84 SPEERVGAFOFSSTPHLEFPIDSFSTQOEVARIKRMVFKGRTETELALXYLHRLGP 143
 QY 121 GGRNASVPOILLIYVDGSGQGVVALPSKOLKRGTVFVGVFRPMBEHLALASPRQ 180
 DB 144 GGRNASVPOILLIYVDGSGQGVVALPSKOLKRGTVFVGVFRPMBEHLALASPRQ 203
 QY 181 HYLAEQVEDATNGLFSTLSSAICSSATPDCEVAHPCEHRTLEWREFFAGNAPCMRGS 240
 DB 204 HYLAEQVEDATNGLFSTLSSAICSSATPDCEVAHPCEHRTLEWREFFAGNAPCMRGS 263
 QY 241 RRTLAIVLAHCPFPYSMKRVEFLTHPATCYRTTCPCPDSPQCONGGTCVPEGIDYQCLCP 300
 DB 264 RRTLAIVLAHCPFPYSMKRVEFLTHPATCYRTTCPCPDSPQCONGGTCVPEGIDYQCLCP 323
 QY 301 IAFGEANALTLSECRVDLLFLDSSAGTLLDGTLPKAVFKSFVRAVLSBDSBARVG 360
 DB 324 IAFGEANALTLSECRVDLLFLDSSAGTLLDGTLPKAVFKSFVRAVLSBDSBARVG 383
 QY 361 VATYSRELLVAVPGEYQDVPDLVMSLDGI PRFGPTLLGSAALRQAAERGFSSATRTQGD 420
 DB 384 VATYSRELLVAVPGEYQDVPDLVMSLDGI PRFGPTLLGSAALRQAAERGFSSATRTQGD 443
 QY 421 RRRRVVLLTSHSEDEVAGPARHARARELLLLGVSEAVRALEBEITGSPKHWVWYSDP 480
 DB 444 RRRRVVLLTSHSEDEVAGPARHARARELLLLGVSEAVRALEBEITGSPKHWVWYSDP 503
 QY 481 QDLFNOIPELOGLKCSRRPGCRTOALDVFMLDTSASVGPENFAOMOSFVMSCALQFEV 540
 DB 504 QDLFNOIPELOGLKCSRRPGCRTOALDVFMLDTSASVGPENFAOMOSFVMSCALQFEV 563
 QY 541 NPDVTOVGLVVGSGVQVTFAGLDTKPTRAMLRATISQAPYLGVSAGTALLHIYDKVMT 600
 DB 564 NPDVTOVGLVVGSGVQVTFAGLDTKPTRAMLRATISQAPYLGVSAGTALLHIYDKVMT 623
 QY 601 VORGARPGVPKAVVVLTTGGRGAEDAAVPAQKLRNNGISVLLVVGVPVLSBGLRRLAGPRD 660
 DB 624 VORGARPGVPKAVVVLTTGGRGAEDAAVPAQKLRNNGISVLLVVGVPVLSBGLRRLAGPRD 683
 QY 661 SLIHVAAYVADLRHYQDVLIEMLCGEAKOPVNLCKPSPCMNEGSCVLTQNGSYRCKCRDGE 720
 DB 684 SLIHVAAYVADLRHYQDVLIEMLCGEAKOPVNLCKPSPCMNEGSCVLTQNGSYRCKCRDGE 743
 QY 721 GPHCENR 727
 DB 744 GPHCENR 750

RESULT 11

ABUS6622
 ID ABUS6622 standard; protein; 807 AA.

AC ABUS6622;

DT 02-APR-2003 (first entry)

DE Lung cancer-associated polypeptide #215.

XX Lung cancer-associated polypeptide; cytostatic; emphysema;

XX antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;

XX small cell lung cancer; benign lesion; precancerous lesion; bronchitis;

XX chronic obstructive pulmonary disease; hyperresistivity pneumonitis;

XX

PT patient, comprises contacting a biological sample from the patient with a
 PR polynucleotide that hybridizes to an ovarian cancer gene.
 XX
 PS Claim 13; Page 304; 332pp; English.

XX The invention relates to a method of detecting an ovarian cancer-
 CC associated transcript in a cell from a patient, by contacting a
 CC biological sample from the patient with a polynucleotide that selectively
 CC hybridizes to a sequence at least 80% identical to any of one of 80
 CC nucleic acid sequences given in the specification. The method is useful
 CC in diagnosing ovarian cancer and in identifying and using agents and/or
 CC targets that inhibit ovarian cancer. The nucleic acid molecule,
 CC polypeptide and the antibody may also be used in detecting ovarian
 CC cancer, monitoring and early detection of relapse following treatment,
 CC monitoring response to therapy, selecting patients for post-operative
 CC chemotherapy or radiation therapy, in selecting mode of therapy,
 CC determining tumour prognosis, early detection of pre-cancerous lesions,
 CC and as vaccines. This sequence corresponds to one of the proteins used
 CC for the detection method of the invention.

XX Sequence 807 AA;

Query Match 99.3%; Score 3799; DB 7; Length 807;
 Best Local Similarity 100.0%; Pred. No. 1.1e-297;
 Matches 727; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQEVHYSKRTIGKISAAKMMCSAAVDIMFLDGSNSYKGSFESKHPATVCGLDI 60
 DB 24 LOEVHYSKRTIGKISAAKMMCSAAVDIMFLDGSNSYKGSFESKHPATVCGLDI 83
 QY 61 SPERVAVGAFQFSSTPHLEPPLDSFSTQOEVKARIKRMVFKGRTETELAKYLHRLGP 120
 DB 84 SPERVAVGAFQFSSTPHLEPPLDSFSTQOEVKARIKRMVFKGRTETELAKYLHRLGP 143
 QY 121 GGRNAAVPOILIIITVDGKSGDVALPSKQIKERGVTVFVAVGVPFPMBEIHALASEPRQ 180
 DB 144 GGRNAAVPOILIIITVDGKSGDVALPSKQIKERGVTVFVAVGVPFPMBEIHALASEPRQ 203
 QY 181 HYLLAAQVEDATNGLTSTLSSAICSSATPDCVNAHPCRHRLTENVREPAAGNAPCWRS 240
 DB 204 HYLLAAQVEDATNGLTSTLSSAICSSATPDCVNAHPCRHRLTENVREPAAGNAPCWRS 263
 QY 241 RRTLAVLAHAHCFPSYMKRVFLTPATCYRTTCGPGCDSPQCGGTCVPEGLDGYCCLP 300
 DB 264 RRTLAVLAHAHCFPSYMKRVFLTPATCYRTTCGPGCDSPQCGGTCVPEGLDGYCCLP 323
 QY 301 LAFGEANCAKLKSLBCRVDLFLDSSAGTTLDFLRAKVFYKRFVRAVLSHDSRAVY 360
 DB 324 LAFGEANCAKLKSLBCRVDLFLDSSAGTTLDFLRAKVFYKRFVRAVLSHDSRAVY 383
 QY 361 VATYSRELLVAVVGEYQVDPDLVMSLDGIPFRGGPTLTGSALROAERFGSATRTGQD 420
 DB 384 VATYSRELLVAVVGEYQVDPDLVMSLDGIPFRGGPTLTGSALROAERFGSATRTGQD 443
 QY 421 RPRRVVLLTSHSEBEVNGPARHARRELLLVGSEARAELEETTSPEKVMVYSQD 480
 DB 444 RPRRVVLLTSHSEBEVNGPARHARRELLLVGSEARAELEETTSPEKVMVYSQD 503
 QY 481 QDFNFIPELQGLKSGRQPGCRTQALDLVFMLDTSASVPEPFAQMSFVRSQALQFV 540
 DB 504 QDFNFIPELQGLKSGRQPGCRTQALDLVFMLDTSASVPEPFAQMSFVRSQALQFV 563
 QY 541 NPVTVQGLVYVSSQVTAFAGLDTKPTRAAMLRAISQAPLVGGVSGAGTLLHITYKWT 600
 DB 564 NPVTVQGLVYVSSQVTAFAGLDTKPTRAAMLRAISQAPLVGGVSGAGTLLHITYKWT 623
 QY 601 VQGGARGVKAVVYLTVGSGADBAVPAQKLRNNGISVYVWGVPVLSGLRLAGPR 660
 DB 624 VQGGARGVKAVVYLTVGSGADBAVPAQKLRNNGISVYVWGVPVLSGLRLAGPR 683
 QY 661 SLIHVAAYADLRHYODVLEMLGSEAKQPNVLCPSQNNGSCVLYONGSYRCKCDGWE 720
 DB 684 SLIHVAAYADLRHYODVLEMLGSEAKQPNVLCPSQNNGSCVLYONGSYRCKCDGWE 743

QY 721 GPHCENR 727
 DB 744 GPHCENR 750

RESULT 13

ADN39869

ID ADN39869 strand1; protein; 807 AA.

ADN39869;

17-JUN-2004 (first entry)

Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:C239.

Human; differential expression: cancer; angiogenic disorder;
 fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 inflammatory disease; autoimmune disease;
 retinal neovascularization syndrome; scarring; uterine fibroid;
 detection; diagnosis; prognosis; drug screening; drug targeting;
 wound healing; contraception; cytostatic; cardiac; immunomodulatory;
 vulnerability; gene therapy; vaccine.

Homo sapiens.

MO2003042661-A2.

22-MAY-2003.

13-NOV-2002; 2002WO-US036810.

13-NOV-2001; 2001US-0350666P.

21-NOV-2001; 2001US-0332464P.

29-NOV-2001; 2001US-034393P.

03-DEC-2001; 2001US-0335394P.

14-DEC-2001; 2001US-0340376P.

08-JAN-2002; 2002US-0347211P.

10-JAN-2002; 2002US-0347349P.

08-FEB-2002; 2002US-0355250P.

13-FEB-2002; 2002US-0356714P.

20-FEB-2002; 2002US-0359077P.

29-MAR-2002; 2002US-036809P.

04-APR-2002; 2002US-0370110P.

12-APR-2002; 2002US-0372245P.

05-JUN-2002; 2002US-0386614P.

16-JUL-2002; 2002US-0396839P.

22-JUL-2002; 2002US-039775P.

22-JUL-2002; 2002US-0397845P.

09-SEP-2002; 2002US-0409450P.

(EOSB-) EOS BIOTECHNOLOGY INC.

Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;

Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;

WPI, 2003-468649/44.

N-PSDB; ADN39862.

Determining the presence or absence of a pathological cell in a patient,
 useful for diagnosing, prognosing or treating cancer, comprises detecting
 a nucleic acid in a biological sample.

Claim 12; SEQ ID NO C239; 1385pp; English.

The invention relates to nucleic acids and proteins (ADN3683-ADN40064)
 whose expression is upregulated or downregulated in specific cancers or
 other diseases such as angiogenic or fibrotic disorders, and to methods
 of determining the presence or absence of a pathological cell in a
 patient by detecting a nucleic acid at least 80% identical to those of
 the invention or by detecting a polypeptide of the invention. The
 invention also relates to expression vectors and host cells comprising a
 nucleic acid of the invention; antibodies which specifically bind a

CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularization syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a polypeptide of the invention.

XX
 SO Sequence 807 AA;

Query Match 99.3%; Score 3799; DB 7; Length 807;
 Best Local Similarity 100.0%; Pred.No.1.1e-297;
 Matches 727; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQEVHVKETIGTIGTISAAKMMKCSAAVDIMFLDGSNSVKGSPERSKHPATVCCGLDI 60
 DB 24 LQEVHVKETIGTIGTISAAKMMKCSAAVDIMFLDGSNSVKGSPERSKHPATVCCGLDI 83
 QY 61 SPERVRVGAFQFSSSTPLBEPPLDSFSTQEVKARIKMWFKGRTETELAKYLHRLGP 120
 DB 84 SPERVRVGAFQFSSSTPLBEPPLDSFSTQEVKARIKMWFKGRTETELAKYLHRLGP 143
 QY 121 GGNNAASVPOLLIIVTDGSKGDVALPSKOLKENGVTYFVAVGPRPMBELHALASPRQ 180
 DB 144 GGNNAASVPOLLIIVTDGSKGDVALPSKOLKENGVTYFVAVGPRPMBELHALASPRQ 203
 QY 181 HVLAEQVEDATNGTLFSTLSSSAICSSATPDCEVBAHPCRHRTLEWREBPAGNAPCMRGS 240
 DB 204 HVLAEQVEDATNGTLFSTLSSSAICSSATPDCEVBAHPCRHRTLEWREBPAGNAPCMRGS 263
 QY 241 RRTLAVALAHCFFYSWKRVELTHPATCYRTTCBPCDOPCONGTCVPEGLDGYOCLCP 300
 DB 264 RRTLAVALAHCFFYSWKRVELTHPATCYRTTCBPCDOPCONGTCVPEGLDGYOCLCP 323
 QY 301 LAAGGEMCNLKLSLECRVDLLFLDSSAGTTIDGFLRAVFRKRYRAVLSDSRAV 360
 DB 324 LAAGGEMCNLKLSLECRVDLLFLDSSAGTTIDGFLRAVFRKRYRAVLSDSRAV 383
 QY 361 VATYSRELLVAVVEGYQVDPDLVMSLDGIPFRGGPTLGSALROAABRGFSATRTGOD 420
 DB 384 VATYSRELLVAVVEGYQVDPDLVMSLDGIPFRGGPTLGSALROAABRGFSATRTGOD 443
 QY 421 RPRRVVLLTSHSEDEVAGPAHARARELLLVGSEAVRALEBITGSPKIMVYSDP 480
 DB 444 RPRRVVLLTSHSEDEVAGPAHARARELLLVGSEAVRALEBITGSPKIMVYSDP 503
 QY 481 QDLFNQIPBELQGLCSRORPGCTOALDLYFMLDTSASVGPENFAPQMSFVASCALQFEV 540
 DB 504 QDLFNQIPBELQGLCSRORPGCTOALDLYFMLDTSASVGPENFAPQMSFVASCALQFEV 563
 QY 541 NPDVTQVGLVYVSQVOTAFGLDTKPTBAAMLRAISOAPLGGVGSAGTALHTYKWT 600
 DB 564 NPDVTQVGLVYVSQVOTAFGLDTKPTBAAMLRAISOAPLGGVGSAGTALHTYKWT 623
 QY 601 VQGGAPVGVKAVVLTGGRGADAAVPAQKLNNGISVLVVGCVLSEGLRRLAGPRD 660
 DB 624 VQGGAPVGVKAVVLTGGRGADAAVPAQKLNNGISVLVVGCVLSEGLRRLAGPRD 683
 QY 661 SLIHVAAYADLRHODVLEMLGGEAKOPNLCBPCKMNEGSCVTLQNSYRCKCDGME 720
 DB 684 SLIHVAAYADLRHODVLEMLGGEAKOPNLCBPCKMNEGSCVTLQNSYRCKCDGME 743
 QY 721 GPHCENR 727
 DB 744 GPHCENR 750

RESULT 14
 ADN38804
 ID ADN38804 standard; protein; 807 AA.
 XX

AC ADN38804;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:122.
 XX
 KW Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularization syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiac; immunomodulatory;
 KW vulnerrary; gene therapy; vaccine.
 XX
 OS Homo sapiens.
 XX
 PN MO2003042661-A2.
 XX
 PD 22-MAY-2003.
 XX
 XX 13-NOV-2002; 2002WO-US036810.
 PF 13-NOV-2001; 2001US-0350666P.
 PR 21-NOV-2001; 2001US-0332464P.
 PR 29-NOV-2001; 2001US-0334393P.
 PR 03-DEC-2001; 2001US-0335394P.
 PR 14-DEC-2001; 2001US-0340376P.
 PR 08-JAN-2002; 2002US-0347211P.
 PR 10-JAN-2002; 2002US-0347349P.
 PR 08-FEB-2002; 2002US-0355250P.
 PR 13-FEB-2002; 2002US-0356714P.
 PR 20-FEB-2002; 2002US-0359077P.
 PR 29-MAR-2002; 2002US-0368809P.
 PR 04-APR-2002; 2002US-0370101P.
 PR 12-APR-2002; 2002US-0372246P.
 PR 05-JUN-2002; 2002US-0386614P.
 PR 16-JUL-2002; 2002US-0396839P.
 PR 22-JUL-2002; 2002US-0397775P.
 PR 22-JUL-2002; 2002US-0397845P.
 PR 09-SEP-2002; 2002US-0409450P.
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
 XX
 DR MPI: 2003-468649/44.
 XX
 PT N-PSDB; ADN38803.
 PT
 XX
 PS Claim 12; SEQ ID NO 122; 1385bp; English.
 XX
 CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularization syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a polypeptide of the invention.

SQ Sequence 807 AA;
 Query Match 99.3%; Score 3799; DB 7; Length 807;
 Best Local Similarity 100.0%; Pred. No. 1.1e-297;
 Matches 727; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LOEVHVSKEITIGKISAAKMMWCSAAVDIMFLDGSNSVGKSPERSKHFATVCGGLDI 60
 DB 24 LOEVHVSKEITIGKISAAKMMWCSAAVDIMFLDGSNSVGKSPERSKHFATVCGGLDI 83
 QY 61 SPERVVGAFOFSGSTPHLEPLDSFSTQGEVKARIKRMVFKGRTETELAKYLLHRGLP 120
 DB 84 SPERVVGAFOFSGSTPHLEPLDSFSTQGEVKARIKRMVFKGRTETELAKYLLHRGLP 143
 QY 121 GGRNVAVPOILLIYVTGKSGQDVALPSKQKENGTVFVAVGVFRPWEELHALASPRQ 180
 DB 144 GGRNVAVPOILLIYVTGKSGQDVALPSKQKENGTVFVAVGVFRPWEELHALASPRQ 203
 QY 181 HYLLAQVEDATNGLFSTLSSSAICSSATPDCVREAHPCBHRTLEHVRERFAGNAPCMRGS 240
 DB 204 HYLLAQVEDATNGLFSTLSSSAICSSATPDCVREAHPCBHRTLEHVRERFAGNAPCMRGS 263
 QY 241 RRTLAVALAHCPEYSKRVFLTHPATCYRTTCPCDSDPCQNGGTCVEBGLDGYQCLCP 300
 DB 264 RRTLAVALAHCPEYSKRVFLTHPATCYRTTCPCDSDPCQNGGTCVEBGLDGYQCLCP 323
 QY 301 LAFGGANCAKLSLECRVDLFLDSSAGTTIDGFLRAKVPYKRVRAVLSDSRAKRG 360
 DB 324 LAFGGANCAKLSLECRVDLFLDSSAGTTIDGFLRAKVPYKRVRAVLSDSRAKRG 383
 QY 361 VATYSRELLVAVVGEGYQDPDLVMSLDGIPFRGSGPTLGSALROAABGFGSATRTGDD 420
 DB 384 VATYSRELLVAVVGEGYQDPDLVMSLDGIPFRGSGPTLGSALROAABGFGSATRTGDD 443
 QY 421 RPRRVVLLTSHSEDEVAAPAHABARRELLLGVSSEAVRALEBITGSPKIVMYSDP 480
 DB 444 RPRRVVLLTSHSEDEVAAPAHABARRELLLGVSSEAVRALEBITGSPKIVMYSDP 503
 QY 481 ODLFNOIPELOGLCSRRORPGCTOALDLVFMLDTSASGVENFAOMOSFVNSCALQFV 540
 DB 504 ODLFNOIPELOGLCSRRORPGCTOALDLVFMLDTSASGVENFAOMOSFVNSCALQFV 563
 QY 541 NPDVTVGLVVGSGVOTAFGLDKTPRAAMLRAISOAPLGGVSGAGTALLHYTKWT 600
 DB 564 NPDVTVGLVVGSGVOTAFGLDKTPRAAMLRAISOAPLGGVSGAGTALLHYTKWT 623
 QY 601 VORGARPVKAVVVLTTGGRGADAAVPAQKLANNGISVLLVGVGVLSGLRRLLAGPRD 660
 DB 624 VORGARPVKAVVVLTTGGRGADAAVPAQKLANNGISVLLVGVGVLSGLRRLLAGPRD 683
 QY 661 SLIHVAAYADLRHYQDVLIEMLCGEAKOPVNLCKSPCMNBSGCVLQNGSYRCKCRDGE 720
 DB 684 SLIHVAAYADLRHYQDVLIEMLCGEAKOPVNLCKSPCMNBSGCVLQNGSYRCKCRDGE 743
 QY 721 GPHCENR 727
 DB 744 GPHCENR 750

RESULT 15
 ADN39446
 ID ADN39446 standard; protein; 807 AA.
 AC ADN39446;
 XX
 XX 17-JUN-2004 (first entry)
 DT
 XX
 DE Cancer/angiogenesis/fibrosis-related polypeptide. SEQ ID NO:A46.
 XX
 XX Human; differential expression; cancer; angiogenic disorder;
 KM fibrotic disorder; psoriasis; ischemia; heart disease; atherosclerosis;
 KM inflammatory disease; autoimmune disease;
 KM retinal neovascularization syndrome; scarring; uterine fibroid;

KM detection; diagnosis; prognosis; drug screening; drug targeting;
 KM wound healing; contraception; cytostatic; cardiant; immunomodulatory;
 KM vulnerary; gene therapy; vaccine.
 OS Homo sapiens.
 XX
 XX MO2003042661-A2.
 XX
 PD 22-MAY-2003.
 XX
 PF 13-NOV-2002; 2002WO-US036810.
 XX
 PR 13-NOV-2001; 2001US-0350666P.
 PR 21-NOV-2001; 2001US-0332464P.
 PR 29-NOV-2001; 2001US-0334393P.
 PR 03-DEC-2001; 2001US-033594P.
 PR 14-DEC-2001; 2001US-0340376P.
 PR 08-JAN-2002; 2002US-0347211P.
 PR 10-JAN-2002; 2002US-0347349P.
 PR 08-FEB-2002; 2002US-0355250P.
 PR 13-FEB-2002; 2002US-0356714P.
 PR 20-FEB-2002; 2002US-0359077P.
 PR 29-MAR-2002; 2002US-036809P.
 PR 04-APR-2002; 2002US-0370110P.
 PR 12-APR-2002; 2002US-0372246P.
 PR 05-JUN-2002; 2002US-0386614P.
 PR 16-JUL-2002; 2002US-0396839P.
 PR 22-JUL-2002; 2002US-0397775P.
 PR 22-JUL-2002; 2002US-0397845P.
 PR 09-SEP-2002; 2002US-0409450P.
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
 XX
 DR WPI: 2003-468649/44.
 DR N-PSDB; ADN39445.
 XX
 PT Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.
 XX
 PS Claim 12; SEQ ID NO A46; 1385bp; English.
 CC
 CC The invention relates to nucleic acid and proteins (ADN38683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularization syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a polypeptide of the invention.

SQ Sequence 807 AA;
 Query Match 99.3%; Score 3799; DB 7; Length 807;
 Best Local Similarity 100.0%; Pred. No. 1.1e-297;
 Matches 727; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LOEVHVSKEITIGKISAAKMMWCSAAVDIMFLDGSNSVGKSPERSKHFATVCGGLDI 60
 DB 24 LOEVHVSKEITIGKISAAKMMWCSAAVDIMFLDGSNSVGKSPERSKHFATVCGGLDI 83

QY 61 SPERVRVGAFOFSSTPHLEFPPLDSFSTQOEVKARIKRMVFKGGRTELELAKYLLHRLGP 120
| | | | |
Db 84 SPERVRVGAFOFSSTPHLEFPPLDSFSTQOEVKARIKRMVFKGGRTELELAKYLLHRLGP 143
| | | | |
QY 121 GGNASVPOLLIIVTDGKSGDVALPSKOLKERGVTVAQVRFPRMEELHALASEPRQ 180
| | | | |
Db 144 GGNASVPOLLIIVTDGKSGDVALPSKOLKERGVTVAQVRFPRMEELHALASEPRQ 203
| | | | |
QY 181 HVLAAQVEDATNGLFSTLSSSAICSSATPDCRVEAHPCHERTLBMVREFAGNAPCMRGS 240
| | | | |
Db 204 HVLAAQVEDATNGLFSTLSSSAICSSATPDCRVEAHPCHERTLBMVREFAGNAPCMRGS 263
| | | | |
QY 241 RRTLAIVLAHCPRYSWKRVFLTHPATCYRTTCGPGCDSPCQNGGTCVPEGLDGYQCLCP 300
| | | | |
Db 264 RRTLAIVLAHCPRYSWKRVFLTHPATCYRTTCGPGCDSPCQNGGTCVPEGLDGYQCLCP 323
| | | | |
QY 301 LARGEANCAIKLSLEGRVDLFLDSSAGTTLDFLRAKVFVKRFVRAVLSEDSRAVG 360
| | | | |
Db 324 LARGEANCAIKLSLEGRVDLFLDSSAGTTLDFLRAKVFVKRFVRAVLSEDSRAVG 383
| | | | |
QY 361 VATYSRELVAVPVEGYQVDPDLVMSLDGI PFRGPTLTGSALROAERGFSGATRTGOD 420
| | | | |
Db 384 VATYSRELVAVPVEGYQVDPDLVMSLDGI PFRGPTLTGSALROAERGFSGATRTGOD 443
| | | | |
QY 421 RPRRVVLLTSHSEDEBVAQPARHARARELLLGVSBAVRALEBITGSPKHVWYSDP 480
| | | | |
Db 444 RPRRVVLLTSHSEDEBVAQPARHARARELLLGVSBAVRALEBITGSPKHVWYSDP 503
| | | | |
QY 481 ODLFNQIPELQKLCGRORPGCRTOALDLVFMLDTSASVGPENFAOMOSFVRSALQFEV 540
| | | | |
Db 504 ODLFNQIPELQKLCGRORPGCRTOALDLVFMLDTSASVGPENFAOMOSFVRSALQFEV 563
| | | | |
QY 541 NPDVTQGLVWYGSQVQTAFGLDTKPTRAAMLRAISOAPYLGVSAGTALLHIYDKWMT 600
| | | | |
Db 564 NPDVTQGLVWYGSQVQTAFGLDTKPTRAAMLRAISOAPYLGVSAGTALLHIYDKWMT 623
| | | | |
QY 601 VQRGARPGVPKAVVLTGGRGABDAAVPAQKLRNNGISVLVGVGPVLSBGLRRLAGPRD 660
| | | | |
Db 624 VQRGARPGVPKAVVLTGGRGABDAAVPAQKLRNNGISVLVGVGPVLSBGLRRLAGPRD 683
| | | | |
QY 661 SLIHVAAYADLRHYHODVLEMLGGEAKOPVNLCKSPCMNEGS CVLQNGSYRCKCRDGMW 720
| | | | |
Db 684 SLIHVAAYADLRHYHODVLEMLGGEAKOPVNLCKSPCMNEGS CVLQNGSYRCKCRDGMW 743
| | | | |
QY 721 GPHCENR 727
| | | | |
Db 744 GPHCENR 750
| | | | |

Search completed: June 21, 2006, 20:07:57
Job time : 206 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2006, 20:08:18 ; Search time 44 Seconds

(without alignment)
1600.697 Million cell updates/sec

Title: US-10-649-591-3

Perfect score: 3826
Sequence: 1 LQEVHVSKEITGKISASAKM.....CKCRDQWGPHECNFLRRP 732

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*

1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	522.5	13.7	3176	2 CGH3A	collagen alpha 3(V
2	513.5	13.4	3137	2 A37797	collagen alpha 3(V
3	480.5	12.6	3124	2 A40020	collagen alpha 1(X
4	421.5	11.0	1857	2 S31312	collagen alpha 1(X
5	421.5	11.0	1888	2 S78476	collagen alpha 1(X
6	407.5	10.7	1747	2 A45374	collagen alpha 1(X
7	399	10.4	2813	1 VMH	von Willebrand fac
8	398	10.4	493	2 A33809	cartilage matrix p
9	390.5	10.2	500	2 S66522	cartilage matrix p
10	386.5	10.1	496	2 A37979	cartilage matrix p
11	305.5	8.0	741	2 T46488	hypothetical prote
12	301	7.9	2944	2 A54849	collagen alpha 1(V
13	290.5	7.6	550	2 T23760	hypothetical prote
14	288.5	7.5	3051	2 S42373	hypothetical prote
15	283	7.4	929	2 S15027	type XII collagen
16	266.5	7.0	567	2 T28797	hypothetical prote
17	266	7.0	843	2 A40970	undulin 1 - human
18	254	6.6	13055	2 T16580	hypothetical prote
19	225.5	5.9	1029	1 S21169	collagen alpha 2(V
20	224	5.9	1022	2 S04111	collagen alpha 2(V
21	216.5	5.7	724	2 A48569	antigen Em100 - B1
22	216.5	5.7	1018	1 CGH2A	collagen alpha 2(V
23	216	5.6	1019	1 A32856	collagen alpha 2(V
24	215	5.6	1025	2 S34839	collagen alpha 1(V
25	209.5	5.5	2098	2 T18397	protein CTRP - mal
26	204.5	5.3	414	2 PS0033	von Willebrand fac
27	198.5	5.2	427	2 G00039	von Willebrand fac
28	192.5	5.0	371	2 S32604	collagen alpha 2(V
29	189.5	5.0	1523	2 T13953	MEGFs protein - ra

30	189	4.9	1028	1 CGH3A	collagen alpha 1(V
31	188	4.9	1151	2 A45226	integrin alpha-1 c
32	184.5	4.8	712	2 A45638	immunodominant m/c
33	178.5	4.7	1180	2 A35854	integrin alpha-1 c
34	174.5	4.6	272	2 A55348	integrin alpha-1 -
35	172.5	4.5	2809	2 T30213	G-cadherin - sea u
36	172	4.5	1320	2 A56136	Jagged protein pre
37	167	4.4	1328	2 T43060	agrin - electric r
38	166.5	4.4	1153	2 S00551	leukocyte surface
39	166.5	4.4	1163	1 RWHUIC	cell surface glyco
40	166	4.3	1955	1 AGCH	agrin precursor -
41	164.5	4.3	1286	2 A88396	protein M01E10.2
42	164	4.3	640	2 T29784	hypothetical prote
43	164	4.3	2318	2 S45306	notch 3 protein -
44	164	4.3	2321	2 S78549	notch3 protein - h
45	162.5	4.2	746	2 S44792	F09G8.8 protein -

ALIGNMENTS

RESULT 1

CGH3A

collagen alpha 3(VI) chain precursor [validated] - human

N:Contains: collagen alpha 3(VI) chain, splice form A9/NI0(-)

C:Species: Homo sapiens (man)

C>Date: 21-Nov-1993 #sequence revision 12-Nov-1999 #text change 16-Aug-2004

C/Accession: A59140, S13679, S24465, A57083, S28776, S00245, C31952, C25848, S26510, S48

R:Chu, M. L.

submitted to GenBank, May 1998

A:Reference number: A59140

A:Accession: A59140

A>Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-3176 <CHU>

A:Cross-references: UNIPROT:P1211; UNIPARC:UPI000012604F; GB:X52022; NID:g3127925; PIDN

R:Chu, M. L.; Zhang, R. Z.; Pan, T.; Stokes, D.; Conway, D.; Kuo, H. J.; Glanville, R.; May

EMBO J. 9, 385-393, 1990

A>Title: Mosaic structure of globular domains in the human type VI collagen alpha-3 chain

1009.

A:Reference number: S13679; MUID:90151612; PMID:1689238

A:Accession: S13679

A:Molecule type: mRNA

A:Residues: 1-30,237-313, 'CWM', 318-322, 'AR', 326-1815, 'FD', 1818-1819, 'ID', 1822-3176 <CH5>

A:Cross-references: UNIPARC:UPI0000173C14; EMBL:X52022; NID:g3127925

A:Accession: S24465

A:Molecule type: protein

A:Residues: 574-585,965-973, 'X', 975-976,1306-1325,1361-1377,1381-1401,1473-1506, 'X', 1508

A:Cross-references: UNIPARC:UPI0000173C15; UNIPARC:UPI0000173C16; UNIPARC:UPI0000173C17,

1962, 'X', 1964-1965,2018-2037,2374-2410,2445-2459,2466-2469, 'X', 2471-2474,2504-2508, 'X',

CIC, UNIPARC:UPI0000173C1D; UNIPARC:UPI0000173C1E; UNIPARC:UPI0000173C1F; UNIPARC:UPI000

10000173C25; UNIPARC:UPI0000173C26; UNIPARC:UPI0000173C27

J. Biol. Chem. 267, 24082-24089, 1992

A>Title: The human type VI collagen gene. mRNA and protein variants of the alpha3 chain

A:Reference number: S28776; MUID:93054780; PMID:1339440

A:Accession: A57083

A:Molecule type: DNA

A:Residues: 310-328 <ZAN>

A:Cross-references: UNIPARC:UPI0000173C28

A:Accession: S28776

A:Molecule type: mRNA

A:Residues: 32-126, 'AK', 129-136, 'L', 138-236 <ZAZ>

A:Cross-references: UNIPARC:UPI0000173C28; GB:S49432; NID:g260296; PIDN:AA824261.1; PID

R:Chu, M. L.; Mann, K.; Deutzmann, R.; Pribula-Conway, D.; Hsu-Chen, C. C.; Bernard, M. P.,

Eur. J. Biochem. 168, 309-317, 1987

A>Title: Characterization of three constituent chains of collagen type VI by peptide se

A:Reference number: S00126; MUID:88029444; PMID:3665927

A:Accession: S00245

A:Molecule type: mRNA, protein

A:Residues: 2024-2046,2092-2156, 'R', 2203-2208, 'X', 2210-2211, 'X', 2213-2227,2228-2251,22314

A:Cross-references: UNIPARC:UPI000016A705; UNIPARC:UPI0000173C29; UNIPARC:UPI0000173C2A,

ID:g1335034

A>Note: the mRNA portion of the sequence corresponds to residues 2092-2157
R.Chu, M.L.; Conway, D.; Pan, T.; Baldwin, C.; Mann, K.; Deutzmann, R.; Tiimpl, R.
J. Biol. Chem. 263, 18601-18606, 1988

A>Title: Amino acid sequence of the triple-helical domain of human collagen type VI.
A.Reference number: A31952; PMID:8906644; PMID:3198591

A:Molecule type: mRNA

A:Residues: 2038-2373 <CH>

A:Cross-references: UNIPARC:UPI0000173C2F; GB:U04211; GB:M20778

A>Note: parts of this sequence were determined by protein sequencing
R.Wall, D.; Mettel, M.G.; Passage, E.; Van Gont, N.; Pribula-Conway, D.; Mann, K.; Deut-

A>Title: Cloning and chromosomal localization of human genes encoding the three chains c
A.Reference number: A29948; PMID:8816104; PMID:3348212

A:Accession: C29848

A:Molecule type: mRNA

A:Residues: 2092-2151 <WEI>

A:Cross-references: UNIPARC:PI000016A720; GB:M27449; NID:929191; PIDN:AAS2057.1; PID

A>Note: part of this sequence was determined by protein sequencing
B.Jander, R.; Rauteberg, J.; Glanville, R.M.

Eur. J. Biochem. 133, 39-46, 1983

A>Title: Further characterization of the three polypeptide chains of bovine and human st
A.Accession: S26510

A:Reference number: 826506; PMID:83209648; PMID:6852033

A:Molecule type: protein

A:Residues: <SALACVAGVG<,UN>

A:Cross-references: UNIPARC:PI0000173C2P

A>Note: this sequence cannot be reliably placed and probably represents the results from
R.Mayer, U.; Poeschl, E.; Mitsch, R.; Specks, U.; Pan, T.C.; Chu, M.L.; Timpl, R.
Eur. J. Biochem. 225, 573-580, 1994

A>Title: Recombinant expression and properties of the Kunitz-type protease-inhibitor mod
A.Reference number: 548709; PMID:95045506; PMID:7525281

A:Accession: S48709

A:Molecule type: mRNA

A:Residues: <NRWIFFLCLAGRADA*,3102-3176 <MAY>
A>Cross-references: UNIPARC:PI0000173C30

A>Note: engineered sequence to allow isolation of the Kunitz-type proteinase inhibitor h
R.Arnoux, B.; Merigault, K.; Saludjian, P.; Norris, F.; Bjorn, S.; Olsen, O., Jr.;
submitted to the Brookhaven Protein Data Bank, August 1994

A:Reference number: A52812; PDB:1KNTT

A:Contents: annotation; X-ray crystallography, 1.6 angstroms, residues 3106-3160

A>Note: engineered sequence expressed in Saccharomyces cerevisiae strain mt-663
C.Comment: The fibronectin type III repeat homology domain may be released during proces

C:Genetics:

A:Gene: GDB:COL6A3

A>Cross-references: GDB:119066; OMIM:120250

A:Map position: 2q37.3-q37.3

C:Complex: type VI collagen is a heterotrimer of one alpha I(VI) chain (see PIR:CGHUIA),
sociations among trimer amino- and carboxyl-terminal domains (with disulfide bonds).

C:function:

A>Description: structural component of extracellular tissue microfibrils associated with
C:Superfamily: collagen VI

C:Keywords: alternative splicing; blocked amino end; cell binding; coiled coil; extracell

F:1-25/Domains: signal sequence #status predicted <SIG>

F:26-3176/Product: collagen alpha 3(VI) chain #status predicted <MAT1>

F:26-2037/Product: amino-terminal nonhelical #status predicted <ANH>

F:26-30,237-3176/Product: collagen alpha 3(VI) chain, splice form A9/N10(-) #status pred

F:327-203/Domains: von Willebrand factor type A repeat homology <WM01>

F:240-405/Domains: von Willebrand factor type A repeat homology <WM02>

F:443-608/Domains: von Willebrand factor type A repeat homology <WM03>

F:637-802/Domains: von Willebrand factor type A repeat homology <VM04>

F:835-999/Domains: von Willebrand factor type A repeat homology <VM05>

F:1027-1191/Domains: von Willebrand factor type A repeat homology <VM06>

F:1231-1394/Domains: von Willebrand factor type A repeat homology <VM07>

F:1434-1599/Domains: von Willebrand factor type A repeat homology <VM08>

F:1637-1802/Domains: von Willebrand factor type A repeat homology <VM09>

F:1836-2005/Domains: von Willebrand factor type A repeat homology <VM10>

F:2038-2373/Region: interrupted helical

F:2040-2042/Region: cell attachment (R-G-D) motif

F:2136-2138/Region: cell attachment (R-G-D) motif

F:2148-2150/Region: cell attachment (R-G-D) motif

F:2154-2156/Region: cell attachment (R-G-D) motif
F:2370-2372/Region: cell attachment (R-G-D) motif
F:2374-3176/Domain: carboxyl-terminal non-helical #status predicted <CNH>
F:2400-2571/Domain: von Willebrand factor type A repeat homology <W11>
F:2617-2800/Domain: von Willebrand factor type A repeat homology <W12>
F:2865-2986/Region: alanine/lysine/proline/chreonine/valine-rich repeats
F:2987-3072/Domain: fibronectin type III repeat homology <FN3>
F:3111-3161/Domain: animal knittz-type proteinase inhibitor homology <BPI>
F:326/Modified site: pyroglutamate carboxylic acid (Gln) (in mature form) #status predicted
F:1108,116,202,251,2079,2231,2558,2677,2861,3036/Binding site: carboxydrate (Aen) (covalent)
F:2087/Disulfide bonds: interchain #status predicted
F:2100,2206,2239,2316,2319/Modified site: 4-hydroxyproline (Pro) #status experimental
F:2103,2209,2212,2322,2337/Binding site: carboxydrate (Lys) (covalent) #status experimental
F:2103,2209,2212,2322,2337/Modified site: 5-hydroxylysine (Lys) #status experimental
F:3111-3161,3120-3144,3136-3157/Disulfide bonds: #status predicted

Query Match 13.7% Score 522.5; DB 2; Length 3176;
Best Local Similarity 25.5% Pred. No. 1,36-28;
Matches 166; Conservative 100; Mismatches 250; Indels 135; Gaps 17;

QY	25	AAVDIMFLDGSNVGKGFERSKHFAITVCDGLDISPERVVGAFQFSSTPHLEPIDS	84
DB	36	AAADIIFLVDSWITIGEHFQLVREFLVDVKSILAVGENDFHFAIVQFNQPHTEFLINT	95
QY	85	FSTQOEVKARIKRVNPFGRGRTETELAKYLL--HRLPGGRNA--STPOLLITTDKSO	140
DB	96	YRTQOEVLSHISNNYSYIGGNTQGGKLEVIQSHLTRAASSRADVDGVPQVIVLTDHSK	155
QY	141	GDVALPSKOLKERVGVFAVGVRPRWEELHALASEPRGOHVLAE--QVEDATNGLFS	197
DB	156	DGLALPSAEELKSAVDVNFALIGVEDADGALKELTASEPLNMHMFLENFTSLHDVGNLS	215
QY	198	TLSSSAICSSATPDRCVEAHPCBHRITLWVREFGAPNAPCRKGRSRTTLAVLAHCPFSWK	257
DB	216	CVHSSV-----SPERAGDETLEKIDTAQ-----	238
QY	258	RVLFTHATCTCRRTCPCPCSPQPCNGSTCVBSLDGYCLCLPLAFGEANCAKLSLBC	317
DB	239	-----DS-----	240
QY	318	RVLDLFLDSSAGTTLDGFLRAKVFVKRFAVLSEDSRVRGVATYSRELLVAVPGEY	377
DB	241	-ADIIIFLIDSSNNTGSNFAVILDFVLNLEKLPICIGQQLRVGVQPSDEPRTMFSLDTY	299
QY	378	QDVEDVLWSLDGIPFRGSGPTLT--GSALRQAERGFSGSATRTGODRPR-----RVVLLTSES	432
DB	300	STKAQVAGVAKALGAFGAGELANTGLALDFVEVNH--TAAQSRVBEVGPQVLVILSAG	356
QY	433	HSEDEVAGPARHARRELLLVGSEAV--PALEBEITGSPGVAVVTSDDPDLFNQIDELQ	491
DB	357	PPSDEIRYGVVALKQASVVFSGIGCAQNASRAELQHITDNN--LVFTVPE--FRSFGDLQ	412
QY	492	GKLCS-----RQR-----PGCRQL-----DLVFLDTSASVGFENFQMOSFVNSCA	535
DB	413	EKLIPYVGAQRIIVLKPPITVQVEVKKRIDVFLVDSSSLGLANFAIRDFIAKVI	472
QY	536	LOFEVNDVTVQGLVNVGSOVOTAFGLDTRPTAAMRLAI--SOAPYLGVGSGAGTALLHI	594
DB	473	QRLIEIGDQLQVNAVQADTVRPEFYFNHTPTKREVITVARKKKPELDGSLALYTSALDFV	532
QY	595	YDKVMTVORG--ARPGVPKAVNVLTGSGADEAIVPAQKLRNNGISVLVVG	643
DB	533	RNNLFTSSAGYRAEGIPKLLVLTGKSLDEISQPAQELKRSSIMAFAG	583

RESULT 2
A37797
collagen alpha 3 (VI) chain precursor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 16-Aug-2004
C:Accession: A37797; A34270; A32674
F:Doñana, R.; Bomalido, P.; Colombari, A.
F:Cell Biol. 111, 2197-2205, 1990


```

Db      24  LQEVHVSKEITIGKISAASKMMWCSAAVDIMFLDSDNSVGKGSFERSKHFAITVCDGLDI 83
Qy      61  SPERVRVGAFOFSSPTEPHLEFPDLSFTQGEVKARIKRMVFKGRTETELAKYLLHRGLP 120
Db      84  SPERVRVGAFOFSSPTEPHLEFPDLSFTQGEVKARIKRMVFKGRTETELAKYLLHRGLP 143
Qy      121 GGNASVPOILLIIVTDGKSQGDVALPSKQIKERGTVFAVGVPFPMWELHALASBPQQ 180
Db      144  GGNASVPOILLIIVTDGKSQGDVALPSKQIKERGTVFAVGVPFPMWELHALASBPQQ 203
Qy      181  HVLAAQVEDATNGLFSTLSSSAICSAATPDCRVEAHPCHEHRTLEWREFAAGNAPCWRS 240
Db      204  HVLAAQVEDATNGLFSTLSSSAICSAATPDCRVEAHPCHEHRTLEWREFAAGNAPCWRS 263
Qy      241  RRTLAVALAACPFSYMKRVFLTPATCYRTTCGPGCDSPQCONGGTCVPEGLDGYOCLCP 300
Db      264  RRTLAVALAACPFSYMKRVFLTPATCYRTTCGPGCDSPQCONGGTCVPEGLDGYOCLCP 323
Qy      301  LAFGEANCAKLKLSLECRVDLFLDSSAGTTLDGFLRAKVPYKRFVRAVLSDBSRAVG 360
Db      324  LAFGEANCAKLKLSLECRVDLFLDSSAGTTLDGFLRAKVPYKRFVRAVLSDBSRAVG 383
Qy      361  VATSRRELVAVVGEYQDVPDLVMSLDGIPIRGPGPTLGSALROAERGFSGATRTGDD 420
Db      384  VATSRRELVAVVGEYQDVPDLVMSLDGIPIRGPGPTLGSALROAERGFSGATRTGDD 443
Qy      421  RPRRVVLLTSHSEDEVAQPARHARARELLILGVSEAVRALEETIGSPKHVMYSDP 480
Db      444  RPRRVVLLTSHSEDEVAQPARHARARELLILGVSEAVRALEETIGSPKHVMYSDP 503
Qy      481  ODLFNOIPELOQKLCRORPGCRTOALDVLVMDLTSASVGPENFAOMOSFVNSCALQFEV 540
Db      504  ODLFNOIPELOQKLCRORPGCRTOALDVLVMDLTSASVGPENFAOMOSFVNSCALQFEV 563
Qy      541  NPDVTQVLVYVGSQVQTAFGDLTKPTRAAMLRAISQAPYLGGVSGAGTALLHIYDKWMT 600
Db      564  NPDVTQVLVYVGSQVQTAFGDLTKPTRAAMLRAISQAPYLGGVSGAGTALLHIYDKWMT 623
Qy      601  VORGARPGVPKAVVVLTTGSGAEDAAVPAQKLRNNGISVLYVVGCVLSEGLRRLAGPRD 660
Db      624  VORGARPGVPKAVVVLTTGSGAEDAAVPAQKLRNNGISVLYVVGCVLSEGLRRLAGPRD 683
Qy      661  SLIHLAAVADLRHQVLIEMLCGEAKOPVNLCKPSPCMNBSGCVLQNGSYRCKCDGME 720
Db      684  SLIHLAAVADLRHQVLIEMLCGEAKOPVNLCKPSPCMNBSGCVLQNGSYRCKCDGME 743
Qy      721  GPHCENRFLRRP 732
Db      744  GPHCENRFLRRP 755

RESULT 2
Q70U28_HUMAN PRELIMINARY; PRT; 755 AA.
AC 070U28;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DE 07-FEB-2006, entry version 11.
DE AMACO precursor.
GN Name=AMACO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_Taxid=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14506275; DOI=10.1074/jbc.M307794200;
RA Sengle G., Koble B., Morgelein M., Paulsson M., Wagener R.;
RT "Identification and characterization of AMACO, a new member of the von
RT Willebrand factor A-like domain protein superfamily with a regulated
RT expression in the kidney.";
J. Biol. Chem. 278:50240-50249(2003).

```

```

CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonCommercial license
CC -----
DR EMBL; AJ536328; CAD60276.1; -, mRNA.
DR GenBank; U0000000165816; Homo sapiens.
DR GO; GO:0008083; P:growth factor activity; IEA.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca-bd.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR013032; EGF_Like_reg.
DR InterPro; IPR02035; VWF_A.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF00092; VMA; 3.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00327; VMA; 3.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS50026; EGF_3; 2.
DR PROSITE; PS50234; VMPA; 3.
DR EGF-like domain; Growth factor; Repeat; Signal.
FT SIGNAL 1
FT CHAIN 117..755 AMACO.
SQ SEQUENCE 755 AA; 82040 MW; 3144AACD34E31A4F CRC64;

Query Match 99.9%; Score 3822; DB 2; Length 755;
Best Local Similarity 99.9%; Pred. No. 1,86-260;
Matches 731; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1  LQEVHVSKEITIGKISAASKMMWCSAAVDIMFLDSDNSVGKGSFERSKHFAITVCDGLDI 60
Db      24  LQEVHVSKEITIGKISAASKMMWCSAAVDIMFLDSDNSVGKGSFERSKHFAITVCDGLDI 83
Qy      61  SPERVRVGAFOFSSPTEPHLEFPDLSFTQGEVKARIKRMVFKGRTETELAKYLLHRGLP 120
Db      84  SPERVRVGAFOFSSPTEPHLEFPDLSFTQGEVKARIKRMVFKGRTETELAKYLLHRGLP 143
Qy      121 GGNASVPOILLIIVTDGKSQGDVALPSKQIKERGTVFAVGVPFPMWELHALASBPQQ 180
Db      144  GGNASVPOILLIIVTDGKSQGDVALPSKQIKERGTVFAVGVPFPMWELHALASBPQQ 203
Qy      181  HVLAAQVEDATNGLFSTLSSSAICSAATPDCRVEAHPCHEHRTLEWREFAAGNAPCWRS 240
Db      204  HVLAAQVEDATNGLFSTLSSSAICSAATPDCRVEAHPCHEHRTLEWREFAAGNAPCWRS 263
Qy      241  RRTLAVALAACPFSYMKRVFLTPATCYRTTCGPGCDSPQCONGGTCVPEGLDGYOCLCP 300
Db      264  RRTLAVALAACPFSYMKRVFLTPATCYRTTCGPGCDSPQCONGGTCVPEGLDGYOCLCP 323
Qy      301  LAFGEANCAKLKLSLECRVDLFLDSSAGTTLDGFLRAKVPYKRFVRAVLSDBSRAVG 360
Db      324  LAFGEANCAKLKLSLECRVDLFLDSSAGTTLDGFLRAKVPYKRFVRAVLSDBSRAVG 383
Qy      361  VATSRRELVAVVGEYQDVPDLVMSLDGIPIRGPGPTLGSALROAERGFSGATRTGDD 420
Db      384  VATSRRELVAVVGEYQDVPDLVMSLDGIPIRGPGPTLGSALROAERGFSGATRTGDD 443
Qy      421  RPRRVVLLTSHSEDEVAQPARHARARELLILGVSEAVRALEETIGSPKHVMYSDP 480
Db      444  RPRRVVLLTSHSEDEVAQPARHARARELLILGVSEAVRALEETIGSPKHVMYSDP 503
Qy      481  ODLFNOIPELOQKLCRORPGCRTOALDVLVMDLTSASVGPENFAOMOSFVNSCALQFEV 540
Db      504  ODLFNOIPELOQKLCRORPGCRTOALDVLVMDLTSASVGPENFAOMOSFVNSCALQFEV 563
Qy      541  NPDVTQVLVYVGSQVQTAFGDLTKPTRAAMLRAISQAPYLGGVSGAGTALLHIYDKWMT 600
Db      564  NPDVTQVLVYVGSQVQTAFGDLTKPTRAAMLRAISQAPYLGGVSGAGTALLHIYDKWMT 623
Qy      601  VORGARPGVPKAVVVLTTGSGAEDAAVPAQKLRNNGISVLYVVGCVLSEGLRRLAGPRD 660

```


Db 323 LAAGEVNCAPKSLBERLDVLFLLSSAGTTLGGFRRAKAFYKRFVQAVLRDSDRAVG 382
Qy 361 VATYSRELIYAVPVGEYQDVPDL 383
Db 383 IASGRILMWAVPC---RGYPAL 402

RESULT 2
US-09-907-794A-34
Sequence 34, Application US/0990794A
Patent No. 6635468
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerltgen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth. J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Thomas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/907,794A
CURRENT FILING DATE: 2001-07-17
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20

; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 34
; LENGTH: 915
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-907-794A-34

Query Match 13.0%; Score 499; DB 2; Length 915;
Best Local Similarity 23.9%; Pred. No. 4e-37;
Matches 188; Conservative 101; Mismatches 264; Indels 232; Gaps 28;

Qy 28 DIMFLDGSNSYKSGKSPERSKHFAITVCDGLDISPERVGAFOFSPHLEFFPLDSFET 87
Db 57 DLVFIIDSSRSVNTHDYAKVEFIVDIIQFLDIGDVTRVGLIGSTYKNEFSLKTFKR 116
Qy 88 QOEVARIKRVVFKGRTETELAKYLLHRLPGGRNA---SVPOLIIVTDKSGQD 142
Db 117 KSEVERAVRMHLSGTMTGLAIQYALINAFSEABGARPLENVPVIMIVTDGRPODS 176
Qy 143 VALPSKOLKEGCVTVFANGVRFPRWEELHALASERGGVTLA---EQVEDATNGLPSTL 199
Db 177 VAEVAKAKADTGILFALGVGVDPNTLKSIGSEPHEDHVFIVANFSQIETLTSVFOKL 236
Qy 200 SSSAICSSATPD-----CRVHAHPCHEHRTLE 225
Db 237 CTAMHCSTLHNCAHFCINIPGSIYCRCKQYIILNSDQTTCHIQDLCAEDHNCQLCVN 296
Qy 226 MYREBAGNAPCWRG-----SRRTLAVLAHCP-----FYSWKRVPFLT 262
Db 297 VPGSFV--CQCYSGYALAEGRCAVAV--DYCASBNHCHECVNADSGYLCCHEGFAL 352
Qy 263 HP--ATCYRTT-C---PGPDSQ-----PCQNGCTVPEGL----- 292
Db 353 NDEKTCRINYCALNKPQ--CEHECVNNEESYCRHGRYTLDPNGKTCRVDHCAQDDH 411
Qy 293 -----DGYOCLCPALF-----GGEANCAKLSLEGVLDL 322
Db 412 GGEQLCLNTEBDSFVQCSEBGLINEDLTKSRVDYCLLSDHCEYSCV-----NMDSR 464
Qy 333 FLDDSSAGTTL--DGLRAKV-----FV--RRFYRAVLSEDSRARV 359
Db 465 FACQCPBGHVLRSDDKTCAKLDSGALGDHGEHSCVSESDSFVCCFGYILRED----- 519
Qy 360 GVATYSRELIYAVPVGEYQDVPDLVMSLDGIPFRGGPTLTGSALRQAAREGSGSATRTGQ 419
Db 520 GATCRKQVCOALIDG-----CEHICVNSDSDYTCCL-----EGF---RLAE 559
Qy 420 D--RPRRVVLLTSHSEDEVA---GPARHARARELLLVGSEAVRALEBITGSPKHV 474
Db 560 DKKRRKDDVCKSTHGCHEHLCVANNNGSYICKSGFVL-----AE----- 600
Qy 475 MYVSDPDLFNOIPELQKLSRQRPGRGTALDLVFMULTSASVGPENFAOMQSFVNSC 534
Db 601 -----DERRCKK-----CTEGPIDLVFYIDSKSGISENFVAVQFVGCI 640
Qy 535 ALQFEVNDVTVQGLVYVGSQVQTRFGUDTPTRAAMLRAISQAPYLGSGSAGTALHI 594
Db 641 IDSLTISPKARVGLQYSTQVHTEFTLRNFSADMKKAAVAHMKYMGKSGMTGIALKGM 700
Qy 595 YDKVATVQGRAP---GVPKAVVVLTVGGRGADEAAVAPQKLRNNGISVLYVGVGVPLSEG 651
Db 701 PERSTQGEGARPLSTVRPRAIVTTDSRADDDVSEWASKAKANGITMYAVGVCAIEBE 760
Qy 652 LRLRLA 656
Db 761 LQELIA 765

RESULT 3

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2006, 20:25:52 ; Search time 23 Seconds

(without alignments)
727.085 Million cell updates/sec

Title: US-10-649-591-3

Perfect score: 3826

Sequence: 1 LQEVNVSKEITGKISASNM.....CKCRDGEHPHCENFLARP 732

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 99297 seqs, 22845552 residues

Total number of hits satisfying chosen parameters: 99297

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA.New:*
1: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US09_NEM_PUB pep.*
2: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US06_NEM_PUB pep.*
3: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US07_NEM_PUB pep.*
4: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US08_NEM_PUB pep.*
5: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US09_NEM_PUB pep.*
6: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US10_NEM_PUB pep.*
7: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US11_NEM_PUB pep.*
8: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US06_NEM_PUB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	324.5	8.5	678	US-10-196-749-150	Sequence 150, App
2	324.5	8.5	678	US-11-101-316-34	Sequence 34, App1
3	285	7.4	550	US-10-196-749-46	Sequence 46, App1
4	242	6.3	180	US-09-970-076-3	Sequence 3, App1
5	210	5.5	769	US-11-165-586-27	Sequence 27, App1
6	189	4.9	436	US-10-505-928-592	Sequence 592, App
7	188	4.9	1151	US-11-246-999-103	Sequence 103, App
8	184.5	4.8	724	US-11-165-586-45	Sequence 45, App1
9	179.5	4.7	765	US-11-165-586-36	Sequence 36, App1
10	178	4.7	2026	US-10-505-928-831	Sequence 831, App1
11	173.5	4.5	1218	US-11-178-724-21	Sequence 21, App1
12	173.5	4.5	1523	US-10-196-749-290	Sequence 290, App
13	164.5	4.3	1167	US-11-175-714-162	Sequence 162, App
14	163.5	4.3	1228	US-11-178-724-22	Sequence 22, App1
15	158.5	4.1	1198	US-09-970-076-4	Sequence 4, App1
16	157.5	4.1	1166	US-11-175-714-168	Sequence 168, App1
17	157.5	4.1	1166	US-11-175-714-166	Sequence 166, App
18	157.5	4.1	1188	US-11-175-714-165	Sequence 165, App
19	157	4.1	737	US-10-196-749-38	Sequence 38, App1
20	157	4.1	737	US-11-296-092-15	Sequence 15, App1
21	157	4.1	737	US-11-296-155-15	Sequence 15, App1
22	155.5	4.1	1034	US-11-246-999-43	Sequence 43, App1
23	155.5	4.1	1141	US-11-175-714-150	Sequence 150, App
24	155.5	4.1	1166	US-11-175-714-148	Sequence 148, App
25	155.5	4.1	1188	US-11-175-714-147	Sequence 147, App

26	155.5	4.1	1189	US-11-246-999-35	Sequence 35, App1
27	153	4.0	1152	US-10-511-937-2444	Sequence 2444, App
28	150.5	3.9	437	US-11-293-697-3386	Sequence 3386, App
29	147.5	3.9	182	US-11-175-714-159	Sequence 159, App
30	145.5	3.8	182	US-11-175-714-151	Sequence 151, App
31	134.5	3.5	247	US-11-330-353-4	Sequence 4, App1
32	130.5	3.4	696	US-10-449-902-41608	Sequence 41608, App
33	128.5	3.4	1170	US-10-511-937-3007	Sequence 3007, App
34	126	3.3	4391	US-11-183-325-56	Sequence 56, App1
35	125.5	3.3	618	US-11-178-724-19	Sequence 19, App1
36	125.5	3.3	1047	US-11-293-697-3852	Sequence 3852, App
37	124.5	3.3	536	US-10-449-902-40167	Sequence 40167, App
38	122.5	3.2	685	US-11-175-714-2	Sequence 2, App1
39	122.5	3.2	685	US-11-175-714-28	Sequence 28, App1
40	122.5	3.2	685	US-11-175-714-32	Sequence 32, App1
41	122.5	3.2	685	US-11-175-714-34	Sequence 34, App1
42	122.5	3.2	685	US-11-175-714-36	Sequence 36, App1
43	122.5	3.2	685	US-11-175-714-38	Sequence 38, App1
44	122.5	3.2	685	US-11-178-724-20	Sequence 20, App1
45	122.5	3.2	830	US-11-175-714-11	Sequence 11, App1

ALIGNMENTS

RESULT 1
US-10-196-749-150
Sequence 150, Application US/10196749
Publication No. US20060094864A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P343081C340
CURRENT APPLICATION NUMBER: US/10/196,749
CURRENT FILING DATE: 2002-07-16
PRIOR APPLICATION NUMBER: 10/052586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 150
LENGTH: 678
TYPE: PRT
ORGANISM: Homo Sapien
US-10-196-749-150

Query Match	8.5%	Score 324.5;	DB 6;	Length 678;
Best Local Similarity	20.9%;	Pred. No. 6.8e-19;		
Matches 100; Conservative	79;	Mismatches 192;	Indels 107;	Gaps 10;

[illegible]

```

US-11-101-316-34
RESULT 2
US-11-101-316-34
: Sequence 34, Application US/11101316
: Publication NO. US20060099657A1
: GENERAL INFORMATION:
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, Christopher J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: ANTIBODIES TO A POLYPEPTIDE ENCODED BY A NUCLEIC ACID
: TITLE OF INVENTION: UNDEREXPRESSED IN MELANOMA
: FILE REFERENCE: P323OR1C17C1
: CURRENT APPLICATION NUMBER: US/11/101,316
: CURRENT FILING DATE: 2005-04-06
: PRIOR APPLICATION NUMBER: 10/065526
: PRIOR FILING DATE: 2002-05-03
: PRIOR APPLICATION NUMBER: 10/066867
: PRIOR FILING DATE: 2001-12-06
: PRIOR APPLICATION NUMBER: PCT/US00/23328
: PRIOR FILING DATE: 2000-08-24
: PRIOR APPLICATION NUMBER: 09/380137
: PRIOR FILING DATE: 1999-08-25
: PRIOR APPLICATION NUMBER: PCT/US99/12252
: PRIOR FILING DATE: 1999-06-02
: PRIOR APPLICATION NUMBER: 60/087759
: PRIOR FILING DATE: 1998-06-02
: NUMBER OF SEQ ID NOS: 170
: SEQ ID NO 34
: LENGTH: 678
: TYPE: PRT
: ORGANISM: Homo Sapien
: US-11-101-316-34

```

Best Local Similarity 20.9%; Pred. No. 6.8e-19;
Matches 100; Conservative 79; Mismatches 192; Indels 107; Gaps 10;

```

OY      27  VDIIMFLDSSNSVNGKSPFRSGHVAITYCDGLDISPERVVRGAOFQSPSTPHLEPFLDPSFS  86
Db      292  IDISFLIDSDTSGIKRRFRITQKOLLADVAQALDIBGAPCLMGVQYQDGNPPTHNLKTHT  3515
OY      87  TOGEVARIKRMVFKGRTETELALKYLLHR--GLPGGRNAPSVPQIILIVYDGSKQSDVA  1444
Db      352  NSRDLTKTALEKITQRGGLSNVRBALSIVTKNPFKSNKGNRSGABNVVWMDGMPDQVE  4111
OY      145  LPRKOLKEGCVIYFVANGVFRPMELHLALASEPRGQHYL--LAEQVEDATNGLFSTLSS  201
Db      412  EASRLARESGINIFPTLTGG-----AAENEKQYVBEENFANKVACRTNGEYSI---  4595
OY      202  SAICSGATPDCVEAHPCHEHRTLEWREBFAGNAPCWRSGRRTLLVLAACHCFYSWKEVFL  2615
Db      460  -----HQSWMGLHNTLOPLY-----KRVCD  4800
OY      262  THPATYRTTCCPQPCDOPCONGGTCEBGLDYQCUCPLAFGSEANCAKLSTECRVDL  3211
Db      481  TDRLASCKT-----CLNS-----ADI  4966
OY      322  LFLDSSAGTTLIDGFLRAKVPYKRVFVAVLSEDSBAPVATYSRELLVANPVGEXQDVP  3811
Db      497  GFPIIDSSSVGNGNFRTYLQFVYTNLTKEFEISDIDTRIGAVQIYYEORLEFGFDPKISSKP  5566
OY      382  DLVWSLDGI.PFGCGPTLTGSALRQAAERGFSGSATRTGQDGRRRVVLLTSHSDEYVAGP  4411
Db      557  DIINAIKRVYSGSGTSGAALNFALQLF-----KSKPNKRKMLIITDGRASDYDVRIP  6121
OY      442  ARHARARELLLGVC--SEAVRALELEITGSP--KHWVWYSDPQDLPNQIPELOKQCS  496
Db      613  ANAAHLKGYITTAIGAMPADEEELVATHPHARDHSPFVDENDLHOYVRILIONCT  670

```

RESULT 3
US-10-196-749-46
; Sequence 46, Application US/10196749
; Publication No. US2006094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zhenli
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 22, 2006, 12:17:24 ; Search time 10858 Seconds

(without alignments)
6466.593 Million cell updates/sec

Title: US-10-649-591-3

Perfect score: 3826
Sequence: 1 LGEVHVSKEITGKISASIKH.....CKCRDGEQPHCNFLRRP 732

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ .p2n.model -DEV=xlh
-Q/abs/ABSMER.epool/US10649591/runac.21062006.161710.8168/app_query.fasta-1
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NOR=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs807
-USER=US10649591.0CEN.1.1.5548.0runac.21062006.161710.8168 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb_env.*
2: gb_pat.*
3: gb_ph.*
4: gb_pl.*
5: gb_dr.*
6: gb_ro.*
7: gb_at.*
8: gb_gy.*
9: gb_un.*
10: gb_vl.*
11: gb_ov.*
12: gb_hcg.*
13: gb_in.*
14: gb_cm.*
15: gb_da.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3826	100.0	3097	5	AY572972 Homo sapi
2	3822	99.9	2647	5	AF536328 Homo sapi
3	3540	92.5	2946	2	CQ842058 Sequence

4	3540	92.5	2946	5	AK122716	AK122716 Homo sapi
5	3528	92.2	2782	2	CQ850594	CQ850594 Sequence
6	3528	92.2	2782	5	AK127756	AK127756 Homo sapi
7	3135.5	82.0	2653	6	MM0536329	AF536329 Mus muscu
8	2194.5	57.4	1857	2	BD209726	BD209726 Compositi
9	2194.5	57.4	1857	2	AR341532	AR341532 Sequence
10	2169.5	56.7	3777	11	BC077945	BC077945 Xenopus
11	1958.5	51.2	1220	2	CQ728687	CQ728687 Sequence
12	1616.5	42.3	52487	5	AY572973	AY572973 Homo sapi
13	1616.5	42.3	121774	5	AC005383	AC005383 Homo sapi
14	1477.5	38.6	154390	6	AC125150	AC125150 Mus muscu
15	1477.5	38.6	176356	6	AC161518	AC161518 Mus muscu
16	1375.5	36.0	235745	12	AC119647	AC119647 Rattus no
17	787	20.6	161750	11	CR352283	CR352283 Zebrafish
18	787	20.6	222348	12	CR931999	CR931999 Danio rer
19	786.5	20.6	194152	11	BX65198	BX65198 Zebrafish
20	786.5	20.6	194444	12	CR932360	CR932360 Danio rer
21	551	14.4	5321	6	AF034136	AF034136 Mus muscu
22	529	13.8	8667	2	CQ720672	CQ720672 Sequence
23	525	13.7	10427	7	BV175219	BV175219 sqm77639
24	525	13.7	10488	2	CS080598	CS080598 Sequence
25	525	13.7	10537	7	BV175220	BV175220 sqm77640
26	525	13.7	10538	2	CQ833988	CQ833988 Sequence
27	525	13.7	10558	2	CQ874659	CQ874659 Sequence
28	525	13.7	10558	2	CQ981613	CQ981613 Sequence
29	525	13.7	10558	2	CS030819	CS030819 Sequence
30	525	13.7	10558	2	CS039771	CS039771 Sequence
31	525	13.7	10558	2	AX393287	AX393287 Sequence
32	525	13.7	10558	2	HS020287	HS020287 H. sapiens R
33	525	13.7	10697	2	CQ412877	CQ412877 Sequence
34	519	13.6	2756	2	AX675291	AX675291 Sequence
35	518.5	13.6	2361	2	CQ716197	CQ716197 Sequence
36	513.5	13.4	10057	11	CH004416	M24282 Chicken alp
37	509.5	13.3	3350	5	BC010444	BC010444 Homo sapi
38	508.5	13.3	3711	2	CQ963664	CQ963664 Sequence
39	507.5	13.3	3711	2	CQ963664	CQ963664 Sequence
40	505	13.2	2943	11	BC063920	BC063920 Xenopus t
41	505	13.2	3373	2	AR085071	AR085071 Sequence
42	504	13.2	3449	2	BD075394	BD075394 Secretory
43	504	13.2	3449	2	BD172254	BD172254 Secreted
44	504	13.2	3449	2	BD172573	BD172573 Secreted
45	504	13.2	3449	2	BD172892	BD172892 Secreted

ALIGNMENTS

RESULT 1
AY572972
LOCUS
DEFINITION
Homo sapiens colon cancer secreted protein-2 (CCSP-2) mRNA,
complete cds.
ACCESSION
AY572972
VERSION
AY572972.1
KEYWORDS
GI:50429311
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

REFERENCE
1 (bases 1 to 3097)
Xin,B., Platzter,P., Fink,S.P., Reese,L., Nosrati,A., Willson,J.K.,
Wilson,K. and Markowitz,S.
TITLE
Colon cancer secreted protein-2 (CCSP-2), a novel candidate
oncogene 24 (4), 724-731 (2005)
JOURNAL
PUBMED
15580307
AUTHORS
Xin,B., Platzter,P., Fink,S.P., Reese,L., Nosrati,A., Willson,J.K.,
Wilson,K. and Markowitz,S.D.
TITLES
Direct Submission
Submitted (13-MAR-2004) Medicine, Case Western Reserve University,
11001 Cedar Ave., Cleveland, OH 44106, USA
FEATURES
Location/Qualifiers

```

source
1. .3097
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="10"
/map="10q25-q26"
1. .3097
/gene="CCSP-2"
89. .2356
/gene="CCSP-2"
/note="colon cancer diagnostic marker"
/codon_start=1
/product="colon cancer secreted protein-2"
/protein_id="AAT77225.1"
/db_xref="GI:50429312"
/translation="MPPFLLEAVCVFLFRVPSPSLPQEHVSKETIGKISAKMM
WCSAADIIMFLDGSNVGKSPERSHRAITVCDGLDIPBVRVGAPOPSSTPHLE
PPLDPSFSTQOEYKARIKMYFKSGRTETELATKYLHRLGGHNSVPQILITVDG
KSGQDVALPSKQLEKGVTFVAVGNRPRHEILALASERGHNVLLAEVEDATNGL
FSTLSSSALCSATPPDCRVBAPCEHRTLEMBREFAGNACMRSGRRITLVAHCPF
YSWKRVELTPATCYRTTCGPDSCOPCGNGTCVPEGIDGYOCLCEPLAFGBANCL
KLSEGRVDLLFLDLSAGTTLDGLRAKVFVRPRAVISEDRAVGATVSRRLV
VAVPGEYODVPLVWSLDGIPRSGPTLTGSLROAABRGSRATRTGDRRARVYV
LUTRESBEDVAPARARAREILLGVSBAVRABAEITGSPKVMVTSDDQDLPN
QIPELQKLCSSRQRPQGCRTQALDVFELDTASVGENFAQMSFVRSCLQPEVND
VTQGLVVGSGVQVTAFGIDTKPTRAMLRALISQAPYLGVGSGATLALHYDKMIV
ORGARPGVPAVVVLTGGRGAEDAAVPAQCLRNNGISVLVVGVPVLSERLALAPR
DSLIIHVAAYDLRYHDVLIEMLCGEAKQPVNLCKPSPCMNBSGCVLQNGSYNCKCRD
GMEGPHCEHFLRFP"

```

ORIGIN

Alignment Scores:

```

Pred. No.: 0 Length: 3097
Score: 3826.00 Matches: 732
Percent Similarity: 100.0% Conservative: 0
Beet Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 5 Gaps: 0

```

US-10-649-591-3 (1-732) x AY572972 (1-3097)

```

QY 1 LeuGInGUaHhSVaLSerLyBGluThrILeGlyVaHLeSerAlaAserLyMet 20
DB 158 CTCcAGGAAGTTCATGTAAGCAAGAAACCATCGGGAAGATTTCAGCTGCACCAAAATG 217
QY 21 MetTrpCySeSerAlaAlaValAspIleMetPheLeuLeuAspGlySerAsnSerValGly 40
DB 218 ATGTGGTCTCGGCTGCAAGTGCATCATCTTTCTGTTAATGGGCTTAAACACGCTCGGG 277
QY 41 LyEGlySerPheGluArgSerLySHIsPheAlaILeThrValCySAspGlyLeuAspIle 60
DB 278 AAAGGAGGCTTTGAAAGGTCCAAGCACTTTGCCATCCAGTCTGTGACGGTTCGACATC 337
QY 61 SerProGluArgValArgValIGlyAlaPheGlnPheSerSerThrProHisLeuGluPhe 80
DB 338 AGCCCGAGAGGGTGCAGAGTGGAGCAATTCAGATTCCAGTTCACCTCTCATTCGGAATTC 397
QY 81 ProLeuAspSerPheSerThrGlnGlnGluValIleArgIleLySArgMetValPhe 100
DB 398 CCTTGATTCAATTTTCAACCCCAACAGAAAGTGAAGGCAGAAATCAAGAGAGATGTTTTT 457
QY 101 LyEGlyGlyArgThrGluThrGluLeuAlaLeuLySerLyLeuLeuHisArgGlyLeuPro 120
DB 458 AAAGGAGGGCGCAGCGAGACGGAACCTTGCTGAATACTCTGCACAGAGGTTGGCT 517
QY 121 GlyGlyArgAsnAlaSerValProGlnIleLeuIleIleValThrAspGlyLySArgGln 140
DB 518 GGAAGGAGAAATCTTCTGTGCCCCAGATCTCATCATGCTCATATGGAGAGTCCAG 577
QY 141 GlyAspValAlaLeuProSerLyGlnLeuLyGlyAlaArgGlyValThrValPheAlaVal 160
DB 578 GGGGATGTGCACCTGCATCCAAACACACGTGAAGAAAGGGGTCTCACTGTGTTGCTGTG 637

```

```

QY 161 GlyValArgPheProArgTrpGluGluLeuHisValAlaLeuAlaSerGluProArgGlyGln 180
DB 638 GGGGTcAGGTTTCCAGGtGGAGGAGAGTGCATGCACTGGCCAGCGAGCTTAAGGGCAG 697
QY 181 HisValLeuLeuAlaGluGlnValIGlyAlaPheAlaThrAsnGlyLeuPheSerThrLeuSer 200
DB 698 CAGTCTGTGGCTTACAGAGGAGGAGAGTGCACCAAGCGGCTTTCAGCACCTCAGC 757
QY 201 SerSerAlaIleCySerSerSerAlaThrProAspCyArgValGlnAlaHisProCySGLu 220
DB 758 AGCTCGGCAATCTGCTCCAGCGCCAGCCAGCACTGCAGAGGTCGAGGCTCACCCCTTGAG 817
QY 221 HisArgThrLeuGluMetValArgGluPheAlaGlyAsnAlaProCySArgTrpArgGlySer 240
DB 818 CACAGAGACCTGGAGATGAGTCCGGAGATTCCGCTGGCAATGCCCATGCTGGAAGAGATCG 877
QY 241 ArgArgThrLeuAlaValLeuAlaAlaHisCySProPheLySerTrpLySArgValPhe 260
DB 878 CGCGGAGACCTTGCGGTGCTGGCTGCACACTGCTCTTACAGCTGGAAAGAGATGTTTC 937
QY 261 LeuThrHisProAlaThrCySArgTrpArgThrCySProGlyProCySAspSerGlnPro 280
DB 938 CTAAACCACCTTGCCACCTGCTACAGAGACCACTGCCAGGCCCTGTGACTGCAGCCC 997
QY 281 CySGLAsnGlyGlyThrCySValProGluGlyLeuAspGlyTySGLnCySLeuCySPro 300
DB 998 TGCcAGAAATGAGAGCACAATGTGTTCCAGAAAGACTGAGCGGCTTACAGTGCCTCTGCCCG 1057
QY 301 LeuAlaPheGlyGlyGluAlaAsnCySAlaLeuLySLeuSerLeuGluLyCySArgValAsp 320
DB 1058 CTGGCTTTGGAGGGAGGCTAACTGTGCTCCGAACTGAAGCTGAATGACAGGCTCGAC 1117
QY 321 LeuLeuPheLeuLeuAspSerSerAlaGlyThrThrLeuAspGlyPheLeuArgAlaLyS 340
DB 1118 CTCCTTCTCTGCTGGAACAGCTCTGGGSCACACCTTGAGCGCTTCTGCGGGCCAAH 1177
QY 341 ValPheValIlyArgPheValArgAlaValLeuSerGlyAspSerArgAlaArgValGly 360
DB 1178 GTCTTGTAAGCGGTTTGTGCGGGCCGTGTGAGGAGAGACTCTCGGGCCCGAGTGGGT 1237
QY 361 ValAlaThrTySerArgGluLeuLeuValAlaValProValIGlyGluTyGlnAspVal 380
DB 1238 GTGGCCACATACAGCAGGAGGAGTGTGTGGGGGTCTGTGGGGAGATCAAGAGATGG 1297
QY 381 ProAspLeuValTrpSerLeuAspGlyIleProPheArgGlyGlyProThrLeuThrGly 400
DB 1298 CTTGACTGTGTGAGAGCTTCATGAGCATCTCTTCCGTGTGGCCCCACCTGACGGGC 1357
QY 401 SerAlaLeuArgGlnAlaAlaGluArgGlyPheGlySerAlaThrArgThrGlyGlnAsp 420
DB 1358 AGTGCCTTCCGGCAGGCGGCAAGCGTGGCTTCGGGAGCGCCACAGACAGGCGAGAC 1417
QY 421 ArgProArgArgValValValLeuLeuThrGluSerHisSerGluAspGluValAlaGly 440
DB 1418 CGGCACTGAGAGTGTGTTTGTCTCACTGAGTCACTCCGAGGATAGGTTGGGGCC 1477
QY 441 ProAlaArgHisAlaArgAlaArgGluLeuLeuLeuGlyValGlySergluAlaVal 460
DB 1478 CAAAGCCGTCAGCAGAGGGCGGAGAGCTGCTCTCTGGGTGAGGCTGAGCGCCG 1537
QY 461 ArgAlaGluLeuGluGluIleThrGlySerProLySHIsValMetValTySerAspPro 480
DB 1538 CGGGCAGAGCTGAGAGAGATCAAGAGGAGCCCAAAGCATGTATGGTCTACTCGAGTCT 1597
QY 481 GlnAspLeuPheAsnGlnIleProGluLeuGlnGlyLySLeuCySAspSerArgGlnAspPro 500
DB 1598 CAGATCTGTTTAAcCAAAATCCCTAGCTGCAGGGGAAACTGTGACGCGGAGCGGCCA 1657
QY 501 GlyCySArgThrGlnAlaLeuAspLeuValPheMetLeuAspThrSerAlaSerValGly 520
DB 1658 GGGTCCGAGACCAAGCCCTGAGACTCGCTTTCATGTTGAGACACTCTGCTCAGTAAAGG 1717
QY 521 ProGluAsnPheAlaGlnMetGlnSerPheValArgSerCySAlaLeuGlnPheGluVal 540

```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 22, 2006, 12:16:14 ; Search time 1146 Seconds
(without alignments)
6680.217 Million cell updates/sec

Title: US-10-649-591-3
Perfect score: 3826
Sequence: 1 LGBVHVSXKTIKIGISASACM.....CKCRDQWEGPHCENPLRRP 732

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5244920 seqs, 3486124231 residues
Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+.p2n.model -DRV=xlh
-Q=/abs/ABSSWEB.epool/US10649591/runac_21062006_161708_8157/app_query.fasta.1
-DB=N.GeneSeq -OPMT=faeap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -MODE=LOCAL
-OUTPRT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs807
-USER=US10649591@CSN_1_1_761@runac_21062006_161708_8157 -MCPD=6 -ICPD=3
-NO MMAP -NEG SCORE=0 -WAIT -DSPELOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N.GeneSeq.8.*
1: GeneSeq19808.*
2: GeneSeq19908.*
3: GeneSeq20008.*
4: GeneSeq20018.*
5: GeneSeq20028.*
6: GeneSeq20038.*
7: GeneSeq20048.*
8: GeneSeq20058.*
9: GeneSeq20068.*
10: GeneSeq20078.*
11: GeneSeq20088.*
12: GeneSeq20098.*
13: GeneSeq20108.*
14: GeneSeq20118.*
15: GeneSeq20128.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3826	100.0	2810	8	ABX76352 Lung Canc
2	3826	100.0	2810	10	ADB80522 Ovarian C
3	3826	100.0	2810	11	ADN38805 Cancer/an

ID	Score	Query Match	Length	DB ID	Description
4	3826	100.0	2810	12	ADL70218
5	3826	100.0	2810	12	ADN59610
6	3826	99.9	2668	11	ADN39651
7	3799	99.3	2424	6	ABQ81552
8	3799	99.3	2424	8	ABX76351
9	3799	99.3	2424	10	ADB80520
10	3799	99.3	2424	11	ADN39652
11	3799	99.3	2424	11	ADN38803
12	3799	99.3	3375	6	ABK92207
13	3799	99.3	3375	11	ADN39445
14	3799	99.3	3375	11	ADN39458
15	3799	99.3	3375	11	ADN39564
16	3799	99.3	3375	12	ADQ70145
17	3540	92.5	2946	12	ADQ63544
18	3528	92.2	2782	13	ADQ78557
19	3387.5	88.5	3485	5	AA815130
20	3382.5	88.4	2481	8	ABX76350
21	3382.5	88.4	2481	10	ADB80518
22	3382.5	88.4	2481	11	ADN38801
23	3178.5	83.1	5464	12	ADG78853
24	3178.5	83.1	5543	12	ADG78852
25	2194.5	57.4	1837	3	AAZ61803
26	2194.5	57.4	1837	4	AAZ61803
27	2194.5	57.4	1837	6	ABJ34888
28	927	24.2	552	12	ADQ21815
29	530.5	13.9	9075	13	ACN41218
30	529	13.7	10542	13	ACN37863
31	525	13.7	10057	13	ADP54236
32	525	13.7	10488	14	ADZ22646
33	525	13.7	10556	4	AA159459
34	525	13.7	10558	6	ABJ92097
35	525	13.7	10558	6	ABX04155
36	525	13.7	10558	8	ABZ76357
37	525	13.7	10558	10	ABX72022
38	525	13.7	10558	11	ADP65610
39	525	13.7	10558	11	ADP65057
40	525	13.7	10558	12	ADQ18246
41	525	13.7	10558	12	ADQ29598
42	525	13.7	10558	13	ADH87615
43	525	13.7	10558	13	ADU06244
44	525	13.7	10558	14	ADV70142
45	525	13.7	10558	14	ADY14519

ALIGNMENTS

RESULT 1
ABX76352
ID ABX76352 standard; DNA; 2810 BP.
XX
AC ABX76352;
XX
DT 02-APR-2003 (first entry)
XX
DB Lung cancer-associated polynucleotide #216.
XX
KW Lung cancer-associated polynucleotide; gene; ds; cytosratic; emphysema;
KW antiinflammatory; antiaesthetic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersecretory pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX
OS Unidentified.
XX
XX
XX WO200286443-A2.
XX
XX 31-OCT-2002.
XX
XX 18-APR-2002; 2002WO-US012476.
XX
XX 18-APR-2001; 2001US-0284770P.
XX 10-MAY-2001; 2001US-0290492P.
XX 09-NOV-2001; 2001US-0339245P.

PR 13-NOV-2001; 2001US-0350666P.
 PR 29-NOV-2001; 2001US-0334370P.
 PR 12-APR-2002; 2002US-0372246P.
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 PI Aziz N, Murray R;
 DR MPI; 2003-093161/08.
 XX P-PSDB; ABUS6623.
 XX
 PT Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer.
 XX
 PS Claim 22; Page 351-352; 453pp; English.
 CC The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
 CC invention
 XX
 SQ Sequence 2810 BP; 554 A; 801 C; 880 G; 575 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 3,93e-244 Length: 2810
 Score: 3826.00 Matches: 732
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0
 US-10-649-591-3 (1-732) x ABX76352 (1-2810)
 QY 1 LeuGInGUvAlhIeVAlSerLySGIuThrIlleGlyVLeIlleSeSerAlaAlaSerLyMet 20
 Db 158 CTCcAGGAAGTCATCTTAAGCAAAAGAAACATCGGGAAGATTTCACCTCCGCAAAAG 217
 QY 21 MetTrpCySserAlaAlaValApIlIeMetPheLeuLeuApGlySerAnsSerValGly 40
 Db 218 ATGTGTGTCTCGGCTCGAGTCAGTCATGTTTCTGTAGATGGGCTTAAAGCGCTCGGG 277
 QY 41 LysGILSerPheGILuRgSerLySerLySHIaPheAlaIlleThrValCyAapGlyLeuApIlIe 60
 Db 278 AAAGGGAGCTTTAAAGGTCCAAGCACTTGGCATCAAGTCGTACAAGGCTCGGACATC 337
 QY 61 SerProGILuRgValaRgValaGlyAlaPheGILnPheserSerThrProHISerGILuPhe 80
 Db 338 AGCCCGAGAGGGCTCAGAGTGGAGCATTCAGTTCCACTCTCATCTCGGAATTC 397
 QY 81 ProLeuApSerPheSerThrGILnGILnGILnValaLysAlaArgIlIeLysArgMetValPhe 100
 Db 398 CCTTGATTCATTTCATCCCAAGCAAGAGTGAAGCAAGATCAAGAGATGGATTTC 457
 QY 101 LysGILGILArgThrGILnRgILnRgILnLeuAlaLeuLyTrLeuLeuHISArgGILyLeuPro 120
 Db 458 AAAGAGGGCGGACGGAAGCGGAACCTTGCTGTAAATACCTTGTGACAGAGGGTTGCT 517

QY 121 GILGILArgAsnAlaSerValProGILnIleLeuIlleValThrAapGILySerGILn 140
 Db 518 GGAGGAGAAATGCTTCTGTGCTGCCAGATCCATCATCATGTCATCATGATGGAAAGTCCAG 577
 QY 141 GILaapValAlaLeuProSerLySGILnLeuLySGILnGlyValaThrValPheAlaVal 160
 Db 578 GGGGATGTGGCACTGCCATCCCAAGCAAGCTGAAGGAAGGGGTGTCACTGTGTTGCTGTG 637
 QY 161 GILValaRgPheProArgTrpGILnGILnLeuHISaIaLeuAlaSerGILuProArgGILn 180
 Db 638 GGGGTCAAGTTTCCAGAGTGGAGAGCTGCATGCATGCGCAGGAGCTTGAAGGCGAG 697
 QY 181 HISaIaLeuLeuAlaGILnGILnValaLysPheAlaThrAnsGILyLeuPheSerThrLeuSer 200
 Db 698 CACGTGCTGTGGCTGAGCAGGTGAGATGTCACCAAGCGGCTCTTCAGCAACCTCAGC 757
 QY 201 SerSerAlaIlleCySserSerAlaThrProAapCyAaRgValaGILnAlaHISProCySGILu 220
 Db 758 ACCTGGCCATCTGCTCCAGCGCCAGCCAGACTGCAGGGTGAAGGCTCACCCCTGTAG 817
 QY 221 HISArgThrLeuGILnMetValaRgILnRgILnGILnAlaProCyAaRgILySer 240
 Db 818 CACAGGACCTGGAATGCTCCGGAGTTCCGTGGCATGCCCATGCTGGAGAGGATG 877
 QY 241 ArgArgThrLeuAlaValaLeuAlaAlaHISCySProPheTrpSerTrpLysArgValaPhe 260
 Db 878 CGGCGGACCTTGCGGTGTGCTGCTGCACTGCTCCCTTCAAGCTGGAAAGAGATGTT 937
 QY 261 LeuThrHISProAlaThrCySArgTrpThrTrpCySProGILyProCyAapSerGILnPro 280
 Db 938 CTAAACCAACCCCTCCAGCTGCTAAGAGACCACTCCAGGCCCTGTGATCTGCGAGGCC 997
 QY 281 CySGILnAnsGILyGILnCySValaProGILnGILnLeuAapGILyTrpGILnCySLeuCySPro 300
 Db 998 TCCAGGAATGAGGAGCATGTGTTCCAGAAAGACTGGAGCGCTTACAGTGCCTTGCCTGG 1057
 QY 301 LeuAlaPheGILyGILnAlaAnsCySAlaLeuLySLeuSerLeuGILnCyAaRgValaAap 320
 Db 1058 CTGGCCTTTGGAGGGAGGCTAACTGTGCCCTGAAGCTGAGCTGGAATGAGAGGTCGAC 1117
 QY 321 LeuLeuPheLeuLeuAapSerSerAlaGILyThrThrLeuAapGILyPheLeuArgAlaLys 340
 Db 1238 GTGGCCACATACAGCAGGAGAGCTGCTGGTGGGGTGCCTGTGGGGAGATGACAGAGTGG 1297
 QY 381 ProAapLeuValaITrPserLeuAapGILyLLeuProPheArgGILyGILyProThrLeuThrGILy 400
 Db 1298 CCTGACCTGTGTGAGGCTCGATGCAATGCTCTTCCGTGTGGGCCCACTTGACGGGGC 1357
 QY 401 SerAlaLeuArgGILnAlaGILnRgILnRgILnPheGILySerAlaThrArgThrGILnGILnAap 420
 Db 1358 AGTGCCTTGGCAGGCGGAGAGAGCTGCTTCCGGAGCGCCACCAAGAGAGGCGCAGAGC 1417
 QY 421 ArgProArgArgValaValaLeuLeuThrGILnSerHISerGILuAapGILnValaAlaGILy 440
 Db 1418 CGGCCACGTAAGAGTGTGGTTTGTCTCACTGAGTCACACTCCGAGGATGAGGTTGGGGGC 1477
 QY 441 ProAlaArgHISaIaArgAlaArgGILnLeuLeuLeuLeuGILnGILySerGILnAlaVala 460
 Db 1478 CCAAGCGGTCAAGCAAGGGCGGAGAGCTGCTCCGTGGGGTGTGGGCAAGTGAAGCCGTG 1537
 QY 461 ArgAlaGILnLeuGILnGILnIlleThrGILySerProLySHISaIaMetValaTrpSerAapPro 480
 Db 1538 CGGCGAGGCTGGAAGAGATCAAGGCAAGCCCAAGCATGTGATGTCTACTCGGATCCT 1597

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 22, 2006, 12:21:53 ; Search time 8668 Seconds
(without alignments)
7085.459 Million cell updates/sec

Title: US-10-649-591-3
Perfect score: 3826
Sequence: 1 LGVHVSKEITKIKISASAKM.....CKCRDQWEGPHCENFLRNP 732

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 48236798 seqs, 27959665780 residues
Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+.p2n.model -DEV=xlh
-Q=/abs/ABSSMB.epool/US10649591/runat_21062006.161713.8205/app_query.fasta_1
-DB=EST -QFMT=faaap -SUFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blonum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs04
-USER=US10649591@CEN_1_1.6323@runat_21062006.161713.8205 -NCPD=6 -ICPD=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST.*
1: gb_ests1.*
2: gb_ests3.*
3: gb_ests4.*
4: gb_ests5.*
5: gb_ests6.*
6: gb_ests7.*
7: gb_ests8.*
8: gb_ests9.*
9: gb_ests10.*
10: gb_ests11.*
11: gb_ests12.*
12: gb_ests13.*
13: gb_ests14.*
14: gb_ests15.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3135.5	82.0	3216	6 AK133486	Mus muscu
2	3130.5	81.8	3039	6 AK145058	Mus muscu
3	3117.5	81.5	2948	6 AK166339	Mus muscu
4	3109.5	81.3	3888	6 AK160971	Mus muscu

5	3109	81.3	3627	6 AK029297	Mus muscu
6	3034.5	79.3	2395	14 D0050604	DO050604 Homo sapi
7	2162	56.5	2261	14 D0050605	DO050605 Pan trogl
8	1234	32.3	896	10 DT808453	DT808453 LB0174.CR
9	1003	26.2	585	9 DA191273	DA191273 DA191273
10	866.5	22.6	124	7 BF164945	BF164945 601778179
11	772.5	20.2	632	11 A2344444	A2344444 1M0078P24
12	706.5	18.5	664	7 BB617396	BB617396 AGENCOURT
13	688	18.0	801	9 CX921133	CX921133 JGI CAAN8
14	677	17.7	577	9 DB146766	DB146766 DB146766
15	654	17.1	469	5 CK392905	CK392905 K0851D11-
16	652	17.0	851	3 BU271772	BU271772 D03506709
17	626	16.4	7 BX956774	7 BX956774 DCF2P781D	CF2P781D AGENCOURT
18	621	16.2	744	5 CF285997	CF285997 AGENCOURT
19	610.5	16.0	638	7 BB613951	BB613951 BB613951
20	556	14.5	804	10 DT201883	DT201883 JGI CAAS1
21	554.5	14.5	572	3 BU262922	BU262922 603816005
22	553.5	14.5	461	9 DN334113	DN334113 LB18185-0
23	541.5	14.2	420	7 BB846685	BB846685 BB846685
24	529	13.8	9534	14 D0052789	D0052789 Homo sapi
25	510	13.3	3955	6 AK054356	AK054356 Mus muscu
26	510	13.3	4728	6 AK087828	AK087828 Mus muscu
27	509.5	13.3	2760	14 AY402756	AY402756 Homo sapi
28	508	13.3	529	9 DA791716	DA791716 DA791716
29	499	13.0	9184	14 AY402246	AY402246 Homo sapi
30	497.5	13.0	647	7 BB655280	BB655280 BB655280
31	493.5	12.9	9534	14 D0052790	D0052790 Pan trogl
32	490	12.8	9022	14 AY402248	AY402248 Mus muscu
33	486.5	12.7	3600	6 CR859355	CR859355 Pongo pyg
34	486	12.7	2760	14 AY402758	AY402758 Mus muscu
35	486	12.7	3549	6 AK138211	AK138211 Mus muscu
36	485	12.7	2993	6 AK163818	AK163818 Mus muscu
37	477	12.5	1746	14 AY417765	AY417765 Homo sapi
38	476.5	12.5	631	7 BB612001	BB612001 BB612001
39	474	12.4	549	8 CN227517	CN227517 RCB020H02
40	468.5	12.2	4126	6 HSM807646	BM647500 Homo sapi
41	468	12.2	2217	6 AK041073	AK041073 Mus muscu
42	467.5	12.2	3180	6 AK028536	AK028536 Mus muscu
43	465	12.2	1752	14 AY417767	AY417767 Mus muscu
44	445	11.6	383	9 DR005740	DR005740 TC11897
45	445	11.6	2760	14 AY402757	AY402757 Pan trogl

ALIGNMENTS

RESULT 1
AK133486
LOCUS
DEFINITION
AK133486 3216 bp mRNA linear HTC 21-Sep-2005
Mus musculus 11 days pregnant adult female ovary and uterus cDNA,
Riken full-length enriched library, clone:5031412M06
Product:VMA-like domains, matrilins and collagens, full insert
sequence.
ACCESSION
AK133486
VERSION
AK133486.1 GI:74225701
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Mus.
REFERENCE
1 Carinci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
10349636
2 Carinci, P. and Hayashizaki, Y.
Normalisation and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
11042159
JOURNAL
PUBMED
REFERENCE

AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komori, H., Akiyama, J., Nishi, K., Kitanishi, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsui, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE
RIKEN integrated sequence analysis (RISA) system-384-format

JOURNAL
Genome Res. 10 (11), 1757-1771 (2000)

PUBMED
11076861

REFERENCE
4

AUTHORS
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Komori, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadoya, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleschmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staudli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Balderelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Guernich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaur, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wymshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohseki, S. and Hayashizaki, Y.

CONSTRM
RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium

TITLE
Functional annotation of a full-length mouse cDNA collection

JOURNAL
Nature 409 (6821), 685-690 (2001)

PUBMED
11217851

REFERENCE
5

AUTHORS
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Oseato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Balderelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Brad, D., Brucic, V., Chochia, C., Corbani, L.E., Cousins, S., Dalia, E., Dignan, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Guernich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedziarski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagaishima, T., Numata, K., Okido, T., Pavan, W.J., Perle, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Raveisi, T., Reed, J.C., Reed, J.J., Reid, J., Ring, B.Z., Ringwald, M., Sanderlin, A., Schneider, C., Semp, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Tesada, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wymshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Komori, H., Nakamura, S., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

CONSTRM
FANTOM Consortium

TITLE
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL
Nature 420 (6915), 563-573 (2002)

PUBMED
12466851

REFERENCE
6

AUTHORS
Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, S., Kodzius, R., Shimokawa, K., Bajic, V.B., Bremner, S.E., Batalov, S., Forrest, A.R.,

Zavolan, M., Davis, M.J., Wilming, L.G., Aldins, V., Allen, J.E., Ambesi-Impolato, A., Apweiler, R., Aurialy, R.N., Bailey, T.L., Banai, M., Baxter, L., Beisel, K.W., Bersano, T., Bono, H., Chalk, A.M., Chiu, K.P., Choudhary, V., Christoffel, A., Clutterbuck, D.R., Crowe, M.L., Dalia, E., Dalrymple, B.P., de Bono, B., Della Gatta, G., di Bernardo, D., Down, T., Engstrom, P., Fagiolini, M., Faulkner, G., Fletcher, C.F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M., Georgii-Hemmen, P., Gingersas, T.R., Gojobori, T., Green, R.E., Guernich, S., Harbers, M., Hayashi, Y., Hensch, T.K., Hirokawa, N., Hill, D., Hummel, L., Iacono, M., Ieko, K., Iwama, A., Ishikawa, T., Jakt, M., Kampin, A., Katoh, M., Kawasawa, Y., Kelso, J., Kitamura, H., Kitano, H., Kollis, G., Krishnan, S.P., Kruger, A., Kummerfeld, S.K., Kurochkin, I.V., Lareau, L.F., Lazarevic, D., Lipovich, L., Liu, J., Liuni, S., McWilliam, S., Madan, B., Madan, M., Mader, M., Marchionni, L., Matsumoto, H., Matsumura, S., Miki, H., Mignone, P., Miyake, S., Morris, K., Motta, Tabar, S., Mulder, N., Nakano, N., Nakachi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, P., Ohara, O., Okazaki, Y., Orlando, V., Pang, K.C., Pavan, W.J., Pavoni, G., Pesole, G., Petrovsky, N., Piazza, S., Reed, J., Reid, J.F., Ring, B.Z., Ringwald, M., Roat, B., Ruan, Y., Salzberg, S.L., Sanderlin, A., Schneider, C., Schonbach, C., Sekiguchi, K., Semp, C.A., Seno, S., Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D., Sinclair, B., Sperling, S., Stupka, E., Sugita, K., Sultana, R., Takenaka, Y., Taki, K., Tammoja, K., Tan, S.L., Tang, S., Taylor, M.S., Tegner, J., Teichmann, S.A., Ueda, H.R., van Nimwegen, B., Verardo, R., Wei, C.L., Yagi, K., Yamashita, H., Zabarovsky, B., Zhu, E., T., Hyde, M., Bult, C., Grimmond, S.M., Tesada, R.D., Liu, E., T., Brucic, V., Quackenbush, J., Wahlestedt, C., Mattick, J.S., Hume, D.A., Kato, K., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori, K., Itoh, M., Suzuki, M., Aoki, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M., Kato, T., Kawai, H., Kawagashira, N., Kawashima, T., Kojima, M., Kondo, S., Komori, H., Nakano, K., Niimura, N., Nishio, T., Okada, M., Plessey, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K., Watanabe, A., Okamura-Ono, Y., Suzuki, H., Kawai, J. and Hayashizaki, Y.

CONSTRM
FANTOM Consortium

TITLE
The transcriptional landscape of the mammalian genome

JOURNAL
Science 309 (5740), 1559-1563 (2005)

PUBMED
16141072

REFERENCE
7

AUTHORS
Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakaniishi, M., Nakamura, M., Nishida, H., Yag, C.C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Raveisi, T., Pang, K.C., Hallinan, J., Mattick, J., Hume, D.A., Lipovich, L., Batalov, S., Engstrom, P.G., Mizuno, Y., Peghini, M.A., Sanderlin, A., Chalk, A.M., Mottagui-Tabar, S., Liang, Z., Lenhard, B. and Wahlestedt, C.

CONSTRM
RIKEN Genome Exploration Research Group

TITLE
Antisense transcription in the mammalian transcriptome

JOURNAL
Science 309 (5740), 1564-1566 (2005)

PUBMED
16141073

REFERENCE
8

AUTHORS
Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hori, F., Iida, J., Imamura, K., Imotani, K., Itoh, M., Kanagawa, S., Kawai, J., Kojima, M., Komori, H., Murata, M., Nakamura, M., Niimura, N., Nishiguchi, H., Nomura, K., Ono, M., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shiraki, T., Tagami, M., Tagami, Y., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y.

CONSTRM
Direct Submission

TITLE
Submitted (30-MAR-2004) Yoshihide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 22, 2006, 12:29:33 ; Search time 349 Seconds

(without alignments)
5886.755 Million cell updates/sec

Title: US-10-649-591-3

Perfect score: 3826

Sequence: 1 LQEVAVSKETICKISASKM.....CKCRDWEQPHCNPLRRP 732

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+pn.model -DEV=xlh
-Q=abs/ABSGMB spool/US10649591/runat 21062006.161715.8251/app_query.fasta -
-DB=abs/ABSGMB spool/US10649591/runat 21062006.161715.8251 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsm62 -TRANS=humand0.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOOCL -OUTFMT=ptc -NORMext -HEAPSIZ=500 -MINLEN=2000000000
-HOST=abs04 -USER=US10649591 @CN 1.1 204 @runat 21062006.161715.8251 -NCPu=6
-ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-NARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YCAPOP=10 -YCAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*

1: /EMC_Celerra_SIDS3/ptodata/2/ina/1_COMB.seq:*
2: /EMC_Celerra_SIDS3/ptodata/2/ina/5_COMB.seq:*
3: /EMC_Celerra_SIDS3/ptodata/2/ina/6A_COMB.seq:*
4: /EMC_Celerra_SIDS3/ptodata/2/ina/6B_COMB.seq:*
5: /EMC_Celerra_SIDS3/ptodata/2/ina/7_COMB.seq:*
6: /EMC_Celerra_SIDS3/ptodata/2/ina/H_COMB.seq:*
7: /EMC_Celerra_SIDS3/ptodata/2/ina/PPCTUS_COMB.seq:*
8: /EMC_Celerra_SIDS3/ptodata/2/ina/PP_COMB.seq:*
9: /EMC_Celerra_SIDS3/ptodata/2/ina/RE_COMB.seq:*
10: /EMC_Celerra_SIDS3/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2194.5	57.4	1837	3	US-09-312-283C-350 Sequence 350, App
2	509.5	13.3	3438	3	US-09-949-016-5648 Sequence 5648, Ap
3	509.5	13.3	3438	3	US-09-949-016-5649 Sequence 5649, Ap
4	505	13.2	3373	2	US-08-897-443-2 Sequence 2, Appl1
5	504	13.2	3449	3	US-09-907-794A-33 Sequence 33, Appl1
6	504	13.2	3449	3	US-09-905-125A-33 Sequence 33, Appl1
7	504	13.2	3449	3	US-09-902-775A-33 Sequence 33, Appl1
8	504	13.2	3449	3	US-09-906-700-33 Sequence 33, Appl1

9	504	13.2	3449	3	US-09-903-603A-33	Sequence 33, Appl1
10	504	13.2	3449	3	US-09-904-920A-33	Sequence 33, Appl1
11	504	13.2	3449	3	US-09-909-064-33	Sequence 33, Appl1
12	504	13.2	3449	3	US-09-905-381A-33	Sequence 33, Appl1
13	504	13.2	3449	3	US-09-906-618-33	Sequence 33, Appl1
14	504	13.2	3449	3	US-09-906-646-33	Sequence 33, Appl1
15	504	13.2	3449	3	US-09-904-462-33	Sequence 33, Appl1
16	504	13.2	3449	3	US-09-902-736A-33	Sequence 33, Appl1
17	504	13.2	3449	3	US-09-906-722A-33	Sequence 33, Appl1
18	504	13.2	3449	4	US-09-905-448-33	Sequence 33, Appl1
19	504	13.2	3449	4	US-09-903-562B-33	Sequence 33, Appl1
20	504	13.2	3449	5	US-09-906-679A-33	Sequence 33, Appl1
21	504	13.2	3449	5	US-09-907-841-33	Sequence 33, Appl1
22	499.5	13.1	3496	3	US-09-949-016-344	Sequence 344, App
23	492	12.9	3447	5	US-10-000-512-9	Sequence 9, Appl1
24	479	12.5	3026	5	US-10-000-512-7	Sequence 7, Appl1
25	477	12.5	2003	3	US-09-949-016-4469	Sequence 4469, Ap
26	477	12.5	2003	3	US-09-949-016-4470	Sequence 4470, Ap
27	451	11.8	3389	3	US-10-104-047-842	Sequence 842, App
28	426	11.1	8802	3	US-08-896-449A-1	Sequence 1, Appl1
29	426	11.1	8802	3	US-09-132-652-1	Sequence 1, Appl1
30	426	11.1	8802	3	US-09-886-900A-1	Sequence 1, Appl1
31	426	11.1	8802	3	US-09-662-478C-1	Sequence 1, Appl1
32	399	10.4	6153	2	US-08-347-594A-1	Sequence 1, Appl1
33	398	10.4	6153	3	US-08-463-682-2	Sequence 2, Appl1
34	396	10.4	8575	3	US-09-381-261A-2	Sequence 5433, Ap
35	386.5	10.1	1807	3	US-09-949-016-5435	Sequence 178, App
36	324.5	8.5	2773	3	US-09-991-181-178	Sequence 178, App
37	324.5	8.5	2773	3	US-09-990-444-178	Sequence 178, App
38	324.5	8.5	2773	3	US-09-990-333-178	Sequence 178, App
39	324.5	8.5	2773	3	US-09-992-598-178	Sequence 178, App
40	324.5	8.5	2773	4	US-09-989-735-178	Sequence 178, App
41	324.5	8.5	2773	5	US-09-989-726-178	Sequence 178, App
42	324.5	8.5	2773	5	US-09-997-528-178	Sequence 178, App
43	324.5	8.5	2773	5	US-09-989-726-178	Sequence 178, App
44	324.5	8.5	2773	5	US-09-997-343-178	Sequence 178, App
45	324.5	8.5	2773	5	US-09-997-653-178	Sequence 178, App

ALIGNMENTS

RESULT 1
US-09-312-283C-350
; Sequence 350, Application US/09312283C
; Patent No. 6573095
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murlson, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312,283C
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 350
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Mouse
US-09-312-283C-350

Alignment Scores:
Pred. No.: 8.93e-200
Score: 2194.50
Percent Similarity: 87.1%
Best Local Similarity: 78.8%
Query Match: 57.4%
DB: 3
Length: 1837
Matches: 434
Conservative: 46
Mismatch: 68
Indels: 7
Gaps: 1

```

US-10-649-591-3 (1-732) x US-09-312-283C-350 (1-1837)
QY      1 LeuGlnGluValHisValSerLySGluThrIleGlyLysIleSerAlaIleSerLysMet 20
      196 CTTCAAGAAATGATGATGAAACCGGAGAACCATGGGAGAAATCGCTGGGCGCAGCAAAATTA 255
QY      21 MetTrpCysSerAlaAlaValAspIleMetPheLeuLeuAspGlySerAsnSerValGly 40
      256 ATGGTGCTCAGCCGGCTGCAATCTGTTTCTTTAGATGGCTTCAACACATCGGG 315
QY      41 LysGlySerPheGlyIuAspSerLysHisPheAlaIleThrValCysAspGlyLeuAspIle 60
      316 AAGGAGACCTTCAGAGAGTCCAAAGCGCTGCGCACTGCTGTGATGCCCTGGAAATC 375
QY      61 SerProGluArgValArgValGlyAlaPheGlnPheSerSerThrProHisLeuGluPhe 80
      376 AGCCTGGCAGAGGTCAGAGTCGAGAGCTTGCAAGTTGGTTCCACTCTCATCTGGAAATTC 435
QY      81 ProLeuAspSerPheSerThrGlnGlnGluValLysAlaArgIleLysArgMetValPhe 100
      436 CCTTGGACTCTTCTCACTCACTGACAGGAATGAAAGCAATCAAGGGGATATGTTTTC 495
QY      101 LysGlyGlyArgThrGluThrGluLeuAlaLeuLysThrLeuLeuHisArgGlyLeuPro 120
      496 AAGAGTGGGCGCACCGAGACGGCGCTAGCCCTAATA--CGCTGAGCAGAGGAGTTTCCC 552
QY      121 GlyIYAArgHisAspValProGlnIleLeuIleIleValThrAspGlyLysSerGln 140
      553 GAGAGCAGAAATGGCTGTGCCCCAGATTCTTATCATCTGCAAGATGGCAAGTCCGAC 612
QY      141 GlyAspValAlaLeuProSerLysGlnLeuLysGluArgGlyValThrValPheAlaVal 160
      613 GGGCCCGTGGCTCTCCCGGCTAAGCAGCTGAGAGAAAGGGGCGATCGTGGTTCGGTA 672
QY      161 GlyValArgPheProArgTrpGluGluLeuHisAlaLeuAspSerLysProArgGlyGln 180
      673 GAGTCCGTTTCCAGGTGGAGCGAGCTGCTACGCTGGCGCACTAGCGGAGGACCG 732
QY      181 HisValLeuLeuAlaGluGlnValGluAspAlaThrAsnGlyLeuPheSerThrLeuSer 200
      733 CATGCTGTGGTGGTGGAGAGTGGAGATGCCAACCAATGGCTCTCCAGACCCCTCAGC 792
QY      201 SerSerAlaIleCysSerSerAlaThrProAspCysArgValGluAlaHisProCysGlu 220
      793 AGCTCCGCACTGCGACCACTGCTGATCCAGACTGAGGGGTGGAACTCATCCCTGTGAG 852
QY      221 HisArgThrLeuGluMetValArgGluPheAlaGlyAsnAlaProCysTrpArgGlySer 240
      853 CGAGAGACGCTGGAGACCGTCAGGAGAGCTGCGGCAATGCCTTGTGTGGAGAGATCA 912
QY      241 ArgArgThrLeuAlaValLeuAlaAlaHisCysProPheTyrSerTrpLysArgValPhe 260
      913 AGGCAAGACAGACACTGTGGTGGCTGCGCTGCTTCTTACAGCTGGAGAGAGATGTC 972
QY      261 LeuThrHisProAlaThrCysTyrArgThrThrCysProGlyProCysAspSerGlnPro 280
      973 CAGACACACCTGCGCAACTGCTACAGAACCATGTGTCCAGGCCCTGTGATCTCCAGCCC 1032
QY      281 CysGlnAsnGlyGlyLysCysValProGluGlyLeuAspGlyLysGlnCysLeuCysPro 300
      1033 TGCAGAAATGAGGCGAGTGCATTCAGAAAGTGTGATAGGATACCACTGTCTTGGCCA 1092
QY      301 LeuAlaPheGlyGlyGluAlaAsnCysValAlaLeuLysLeuSerLeuGluCysArgValAsp 320
      1093 CTGGGACTTGGAGGGGAGATCACTGTGCCCGAACTAGCTGGAATGCAAGATCGAT 1152
QY      321 LeuLeuPheLeuLeuAspSerSerAlaGlyThrThrLeuAspGlyPheLeuArgAlaLys 340
      1153 GTCCCTCTTCTGTGAGACATTCGTGAGGACCACTTGGGGGCTTCCGAGGGGCGCAAG 1212
QY      341 ValPheValLysArgPheValArgAlaValLeuSerGluAspSerArgAlaArgValGly 360
      1213 GCCTTGTCAAGCGCTTGTGAGGCGGTGTGAGGAGAGAACTCCGAGCCGCGCTTGGG 1272

```

```

QY      361 ValAlaThrTyrSerArgGluLeuLeuValAlaVal-ProValGlyGluTyrGlnAsp-V 380
      1273 ATAGCAGATTATGGCAGGAATCTAATAGTGGCGGTCCCTGTGCGGAGATCAAGATTC 1332
QY      380 aLProAspLeuValTrpSerLeuAspGlyIleProPheArgGlyIleProThrLeuThrG 400
      1333 TGCCTGACCTGATCAGAGACCTTGAACAGATTCCTTTCAGCGGTGGCCGACCTTAACG 1392
QY      400 LysPheAlaLeuArgGlnAlaGluArgGlyPheGlySerAlaThrArgThrGlnAla 420
      1393 GAGTGCCTTGTCTCCAGTGGCAGAGCAGCGCTTGGAGTGCAGAGAGACTGTGCAAG 1452
QY      420 SPArgProArgArgValValIleLeuLeuThrGlnSerHisSerGluAspGluValAlaG 440
      1453 ACAGGCGCAGCAGAGATAGTATGTTCTCTCATGMAATCAGCTCCAGATGAGTGTCTG 1512
QY      440 LysProAlaArgHisAlaArgAlaArgGluLeuLeuLeuGlyValGlySerGluAlaV 460
      1513 GGCACAGACTCACCGAAGGCTCGGAGACTACTCTCTGGGCGTGGCAGATGAGATCC 1572
QY      460 aLArgAlaGluLeuGluGluIleThrGlySerProLysHisValMetValTyrSerAsp 480
      1573 TCAAGCGGAGCTGTGAAGATCACCGGTAGCCGAAAGCATGTGATGTTCACACAGAAC 1632
QY      480 roGlnAspLeuPheAsnGlnIleProGluLeuGlnGlyLysLeuCysSerArgGlnArgP 500
      1633 CTCAGAGACTGTTC-AGCCAAAT-CCAGAGCTCAGAGAGGCTATTCAGACCCAGCCGCG 1690
QY      500 roGlyCysArgThrGlnAlaLeuAspLeuValPheMetLeuAspThrSerAlaSerValG 520
      1691 CAGGCTGCCAGCAGACAGTCACTGAGCTGTCTTCTGTGG-ATGCCCTGTGC-TGTGTGG 1748
QY      520 LysProGluAsnPheAlaGlnMetGlnSerPheValArgSerCysValAlaLeuGlnPheGluV 540
      1749 GACGTGAGAACTTGCCTCAAAATGACAGAGCTTATCATGAGAAATGACCCCTCCGGTTTGATG 1808
QY      540 aLAsnProAspValThrGlnValGlyLeu 549
      1809 TGAATCTGATGTGACACAAAGTTGGCTG 1837
Db

```

RESULT 2
 US-09-949-016-5648
 / Sequence 5648, Application US/09949016
 / Patent No. 6812339
 / GENERAL INFORMATION:
 / APPLICANT: VENTER, J. Craig et al.
 / TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 / WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 / FILE REFERENCE: CL001307
 / CURRENT APPLICATION NUMBER: US/09/949,016
 / PRIOR FILING DATE: 2000-04-14
 / PRIOR APPLICATION NUMBER: 60/241,755
 / PRIOR FILING DATE: 2000-10-20
 / PRIOR APPLICATION NUMBER: 60/237,768
 / PRIOR FILING DATE: 2000-10-03
 / PRIOR APPLICATION NUMBER: 60/231,498
 / PRIOR FILING DATE: 2000-09-08
 / NUMBER OF SEQ ID NOS: 207012
 / SOFTWARE: FastSeq for Windows Version 4.0
 / SEQ ID NO 5648
 / LENGTH: 3438
 / TYPE: DNA
 / ORGANISM: Human
 US-09-949-016-5648

Alignment Scores:
 Pred. No.: 4,47e-38 Length: 3438
 Score: 509.50 Matches: 200
 Percent Similarity: 38.3% Conservative: 126
 Best Local Similarity: 23.5% Mismatches: 316
 Query Match: 13.3% Indels: 210
 DB: 3 Gaps: 29

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 22, 2006, 13:14:36 ; Search time 2140 Seconds
(Without alignment)
6304.590 Million cell updates/sec

Title: US-10-649-591-3

Perfect score: 3826
Sequence: 1 LEBVHVSKEITGKISASIKM.....CKCRDQWEGPHCENTRLRRP 732

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-O=/abs/ABSSMB.spool/US10649591/runat_21062006_161725_8442/app_query.fasta_1
-DB=Published Applications NA Main -OPMT=fastlap -SUFFIX=p2n.rnpsm
-KINMATCH=0.1 -LOOPT=0 -UNITS=bits -START=1 -END=1
-MATRIX=blowum62 -TRANS=human40.cdi -LIST=45 -DOCCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ext -NORM=ext
-HEAPSIZE=500 -MNLN=0 -MAXLEN=200000000 -HOST=ab8808
-USER=US10649591 @CGN 1.1 1675 @runat_21062006_161725_8442 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-NARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA Main:

1: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US07_PUBCOMB.seq:*
2: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US08_PUBCOMB.seq:*
3: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US09_PUBCOMB.seq:*
4: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US09C_PUBCOMB.seq:*
5: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US10_PUBCOMB.seq:*
6: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US10C_PUBCOMB.seq:*
7: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US10C_PUBCOMB.seq:*
8: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US10C_PUBCOMB.seq:*
9: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US10C_PUBCOMB.seq:*
10: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US10F_PUBCOMB.seq:*
11: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US10F_PUBCOMB.seq:*
12: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US11_PUBCOMB.seq:*
13: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US11A_PUBCOMB.seq:*
14: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US11B_PUBCOMB.seq:*
15: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US11C_PUBCOMB.seq:*
16: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US11D_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	3826	100.0	2810 7	US-10-295-027-123 Sequence 123, App

2	3826	100.0	2810 7	US-10-173-999-86 Sequence 86, App1
3	3826	100.0	2810 8	US-10-229-345-5 Sequence 5, App1
4	3826	100.0	2810 8	US-10-274-177-5 Sequence 5, App1
5	3826	100.0	2810 8	US-10-650-112-5 Sequence 5, App1
6	3826	100.0	2810 11	US-10-649-591-5 Sequence 5, App1
7	3826	100.0	2810 13	US-11-041-788-5 Sequence 969, App
8	3822	99.9	2268 7	US-10-295-027-969 Sequence 121, App
9	3799	99.3	2424 7	US-10-295-027-121 Sequence 970, App
10	3799	99.3	2424 7	US-10-295-027-970 Sequence 22, App1
11	3799	99.3	2424 7	US-10-087-080-22 Sequence 84, App1
12	3799	99.3	2424 7	US-10-173-999-84 Sequence 1, App1
13	3799	99.3	3375 3	US-09-930-020A-151 Sequence 763, App
14	3799	99.3	3375 7	US-10-295-027-763 Sequence 776, App
15	3799	99.3	3375 7	US-10-295-027-776 Sequence 882, App
16	3799	99.3	3375 9	US-10-295-027-882 Sequence 1, App1
17	3799	99.3	3375 9	US-10-702-180-1 Sequence 1, App1
18	3887.5	88.5	3485 10	US-10-450-763-17334 Sequence 11734, A
19	3382.5	88.4	2481 7	US-10-295-027-119 Sequence 119, App
20	3382.5	88.4	2481 7	US-10-173-999-82 Sequence 82, App1
21	2194.5	57.4	1837 3	US-09-866-050A-350 Sequence 350, App
22	2194.5	57.4	1837 6	US-10-152-661-350 Sequence 350, App
23	927	24.2	352 9	US-10-723-860-4635 Sequence 4635, App
24	857	22.4	1992 6	US-10-027-632-97356 Sequence 97356, A
25	857	22.4	1992 6	US-10-027-632-97356 Sequence 3202, App
26	762	19.9	600 13	US-11-060-756-3202 Sequence 7473, App
27	762	19.9	600 13	US-11-060-756-3202 Sequence 7474, App
28	762	19.9	600 13	US-11-060-756-7474 Sequence 7474, App
29	762	19.9	600 13	US-09-918-715-217 Sequence 217, App
30	525	13.7	10558 7	US-10-096-534-14 Sequence 14, App1
31	525	13.7	10558 8	US-10-372-683-19 Sequence 19, App1
32	525	13.7	10558 8	US-10-734-564-24 Sequence 24, App1
33	525	13.7	10558 9	US-10-723-860-1064 Sequence 1064, App
34	525	13.7	10558 9	US-10-486-678-19 Sequence 19, App1
35	525	13.7	10558 10	US-10-956-157-1237 Sequence 1237, App
36	525	13.7	10558 10	US-10-956-157-1237 Sequence 217, App
37	525	13.7	10558 10	US-10-287-436A-151 Sequence 151, App
38	525	13.7	10558 10	US-10-287-436A-151 Sequence 764, App
39	525	13.7	10558 10	US-10-287-436A-151 Sequence 1948, A
40	525	13.7	10558 10	US-10-287-436A-151 Sequence 43, App1
41	525	13.7	10558 10	US-10-287-436A-151 Sequence 491, App
42	525	13.7	10558 10	US-10-287-436A-151 Sequence 493, App
43	525	13.6	10250 9	US-10-741-600-491 Sequence 491, App
44	521	13.6	10250 9	US-10-741-600-491 Sequence 493, App
45	521	13.6	10944 9	US-10-741-600-493 Sequence 493, App

ALIGNMENTS

RESULT 1
US-10-295-027-123
Sequence 123, Application US/10295027
Publication No. US20030232350A1
GENERAL INFORMATION:
APPLICANT: Afar, Daniel
APPLICANT: Aiz, Natascha
APPLICANT: Ginsberg, Wendy M.
APPLICANT: Gish, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Heyezl, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
APPLICANT: Bob Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
FILE REFERENCE: 018501-012500US
CURRENT APPLICATION NUMBER: US/10/295,027
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394

```

/ PRIOR FILING DATE: 2001-11-15
/ PRIOR APPLICATION NUMBER: US 60/332,464
/ PRIOR FILING DATE: 2001-11-21
/ PRIOR APPLICATION NUMBER: US 60/334,393
/ PRIOR FILING DATE: 2001-11-29
/ PRIOR APPLICATION NUMBER: US 60/340,376
/ PRIOR FILING DATE: 2001-12-14
/ PRIOR APPLICATION NUMBER: US 60/347,211
/ PRIOR FILING DATE: 2002-01-08
/ PRIOR APPLICATION NUMBER: US 60/347,349
/ PRIOR FILING DATE: 2002-01-10
/ PRIOR APPLICATION NUMBER: US 60/355,250
/ PRIOR FILING DATE: 2002-02-08
/ PRIOR APPLICATION NUMBER: US 60/356,714
/ PRIOR FILING DATE: 2002-02-13
/ Remaining Prior Application data removed - See File Wrapper or PAM.
/ NUMBER OF SEQ ID NOS: 1386
/ SOFTWARE: Patent Ver. 2.1
/ SEQ ID NO 123
/ LENGTH: 2810
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-295-027-123

```

```

Alignment Scores:
Pred. No.: 0 Length: 2810
Score: 3826.00 Matches: 732
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 7 Gaps: 0

```

US-10-649-591-3 (1-732) x US-10-295-027-123 (1-2810)

```

QY 1 LeuGluGluValHisValSerIySGluThrIleGlyValIleSerAlaIleSerIyMet 20
DB 158 CTCGAGGAAGTCCATGTACAAAGAAACCATCGGAAAGATTTCAGCTGCAGCAAAATG 217
QY 21 MetTrpCysSerAlaAlaValAspIleMetPheLeuLeuAspGlySerAsnSerValGly 40
DB 218 ATGTGGTCTCGGTGAGTGGACATCATGTTCTGTAAATGGGTTCACAGCGTGGG 277
QY 41 LyeGlySerPheGluArgSerIyHisPheAlaIleThrValCysAspGlyLeuAspIle 60
DB 278 AAAGGGGAGCTTGAAGAGTCCAAAGCACTTGCCATCAGCTGTGACGGTCTGACATC 337
QY 61 SerProGluArgValArgValGlyAlaPheGlnPheSerSerThrProHisLeuGluPhe 80
DB 338 AGCCCCAGAGGGTCAGAGTGGGAGCATTCAGTTCAGTTCCTCTCATCTGGAATTC 397
QY 81 ProLeuAspSerPheSerThrGlnGlnGluValIleValIleValIleValIleValIle 100
DB 398 CCTTGGATTCAATTTTCAACCAAGGAAGTGAAGCAAGATCAAGAGAGATGTTTC 457
QY 101 LyeGlyValArgThrGluThrGluLeuAlaLeuIySerIyLeuLeuHisArgGlyLeuPro 120
DB 458 AAAGGAGGGCGCAGCGAGCGAACTTGCTGAAATACCTTCGACAGAGGGTTCCT 517
QY 121 GlnGlyArgAsnAlaSerValProGlnIleLeuIleIleValIleThrAspGlyIySerGln 140
DB 518 GAGAGCGAAGATCTTGTGCCCCAGATCTTCATCTCATCTGATGGAGATGCCAG 577
QY 141 GlnAspValAlaLeuProSerIyGlnLeuIyGlnArgGlyValIleThrValPheAlaVal 160
DB 578 GGGGATGTGCACTGCATCCAGCACTGAAGAAAGGGGTGTCACTGTGTTGCTGTG 637
QY 161 GlnValArgPheProArgTrpGluGlnLeuHisAlaLeuAlaSerGluProArgGlyGln 180
DB 638 GGGGTGAGTTTCCAGAGTGGAGAGCATGTCATGCTGCAGCGAGCTTAAGAGGCGAG 697
QY 181 HisValLeuLeuAlaGluGlnValGlnAspAlaThrAsnGlyLeuPheSerThrLeuSer 200
DB 698 CAGTGTCTTGTGCTGAGCAGGTGAAGATGCCAACAAGGCTCTTCAAGACCTTCAGC 757

```

```

QY 201 SerSerAlaIleCysSerSerAlaThrProAspCysArgValGluAlaHisProCysGlu 220
DB 758 AGCTGGCCCATCTGCTCCAGCGCCAGCCAGATCGAGGTTCCAGGCTCACCTGTGAG 817
QY 221 HisArgThrLeuGluMetValArgGluPheAlaGlyAsnAlaProCysTrpArgGlySer 240
DB 818 CACAGGACGCTGGAGATGTGCTCCGGAGTTGCTGGCAATCCCCCATCTGAGAGAGATCG 877
QY 241 ArgArgThrLeuAlaValLeuAlaAlaHisCysProPheTrpSerTrpIyAspValPhe 260
DB 878 CGGCGGACCTTCGGGTGCTGGCTGACATGTCCTTCTACAGCTGGAAGAGATGTC 937
QY 261 LeuThrHisProAlaThrCysIyArgThrThrCysProGlyProCysAspSerGlnPro 280
DB 938 CTAAACCACTCTGCACACTGCTCAAGAACCACTGCCAGGCCCTGTGACTGCCAGCCC 997
QY 281 CysGlnAsnGlyGlyThrCysValProGluGlyLeuAspGlyIyTrgIyCysLeuCysPro 300
DB 998 TGCACAGATGAGAGGCATGTGTTCCAGAAAGACTGACCGGCTAACAGTCCCTGCCCCG 1057
QY 301 LeuAlaPheGlyGlyGluAlaAsnCysAlaLeuIyLeuSerIyGluCysArgValAsp 320
DB 1058 CTGGCTTTGGAGGAGGCTAACCTGTGCTTGAAGCTGAGCTTGAAATCAGAGGTGAC 1117
QY 321 LeuLeuPheLeuLeuAspSerSerAlaGlyThrThrLeuAspGlyPheLeuArgAlaIyS 340
DB 1118 CTCCTCTTCTGCTGACAGCTCTGGGGGACCACTGGACGGCTTCTCGGGGCCAAA 1177
QY 341 ValPheValIyAspArgPheValArgAlaValLeuSerGluAspSerArgAlaArgValGly 360
DB 1178 GTCTTGATGAAGCGTTTGTGGCGGCGGTGCTGAGGAGGACTCTCGGGCCCAAGTGGGT 1237
QY 361 ValAlaThrTrpSerArgGluLeuLeuValAlaValProValGlyGluTrgIyAspVal 380
DB 1238 GTGGCACAATGACAGAGAGCTGTGTGGCCGTGCTGTGGGAGAGTACAGAGATGTG 1297
QY 381 ProAspLeuValTrpSerLeuAspGlyIleProPheArgGlyGlyProThrLeuThrGly 400
DB 1298 CTGACCTGTGCTGGAGGCTCGATGACATTCCTTCCGTGTGTGGCCCACTGACGGGC 1357
QY 401 SerAlaLeuArgGlnAlaAlaGluArgGlyPheGlySerAlaThrArgThrGlyGlnAsp 420
DB 1358 AGTGCTTGTGGCGAGCGGAGAGCTGTGGCTTGGGAGCCCAACAGAGAGCGACAGAG 1417
QY 421 ArgProArgArgValValIleLeuLeuThrGlnSerHisSerGluAspGluValAlaGly 440
DB 1418 CGGCACAGTAGAGTGTGTTGTCTCATGAGTCACTCCAGAGATGAGTGTGCGGGC 1477
QY 441 ProAlaArgHisAlaArgAlaArgGluLeuLeuLeuGlyValGlySerGluAlaVal 460
DB 1478 CAGCGCGTCAACGCAAGGGGCGAGAGCTGCTCTGTGGGTGTAAAGCAATGAGCGCTG 1537
QY 461 ArgAlaGluLeuGluGluIleThrGlySerProIyHisValMetValIySerAspPro 480
DB 1538 CGGCGAGAGGTGAGAGATCAACAGCAGCCCAACATGTGTGCTACTCGGATCTCT 1597
QY 481 GlnAspLeuPheAsnGlnIleProGluLeuGlnGlyIyLeuCysSerArgGlnArgPro 500
DB 1598 CAGGATCTGTTCAACCAAAATCCTTGAGCTGCAAGGAGAGCTGTGCAAGCCGAGCGGCA 1657
QY 501 GlnCysArgThrGlnAlaLeuAspLeuValPheMetLeuAspThrSerAlaSerValGly 520
DB 1658 GGGTGCAGGACCAACACCTCGAGCTCGTCTTCAATGTGGAGCACTGTGCTCAGTATGGG 1717
QY 521 ProGluAsnPheAlaGlnMetGlnSerPheValArgSerCysAlaLeuGlnPheGluVal 540
DB 1718 CCCGAGAAATTTGTCTCAGATGCAAGACTTTGTGAGAAAGCTGTGCTCCAGTTTGAGGTG 1777
QY 541 AsnProAspValIleThrGlnValGlyLeuValValIyTrgIySerGlnValGlnThrAlaPhe 560
DB 1778 AACCTGAGGTGACACAGGTGGCTGTGTGTATGTGACGACGAGTGTGACACTGCTTC 1837

```


Db 3658 TACTGACAGAGGCTACTGACACTGCTCTCTCTGATGGCTCTCCAGGCTGTCCAG 3717
 Qy 42 Glycerheg1uargser1ywh1aPheal1a1eThrValCyuaerG1yLeuAer1leSer 61
 Db 3718 GCTGAGTTTGAGTGCCTGAGAGGCTTTGTGTGGACATGATGAGCGGCTGGCATCTCC 3777
 Qy 62 ProG1uargValaTargVala1aPheGlnPheSerSerThrProH1aLeuGluPhePro 81
 Db 3778 CAGAAAGTGGCTCCGCTGGCCGCTGGTGGAGTACACAGACGAGCTCCACGCTCATGCGGG 3837
 Qy 82 LeuAerSerPheSerThrGlnGlnGluVala1a1aTarg1lelyuAerG1yLeuPhe 101
 Db 3838 CTCAAGACCCGGAAGCAGCCTGTCAAGCTGGCGCGCTTGGCAGCCAGTGAAGTACCG 3897
 Qy 102 G1yG1yArg---ThrGluThrGluLeuAlaLeu1yTy1eLeuH1aArgG1yLeuPro 120
 Db 3898 GGCAGCCAGGTGGCTCCACAGAGAGGCTTGAATACACACTGTTCGAA----- 3948
 Qy 121 G1yG1yArgAen1aSerValProGln1leu1le1Val1ThrArgG1yLysSerGln 140
 Db 3949 -----ATCATACAGCAAGATCGACCGCTGAAAGCC 3978
 Qy 141 G1yAerVala1aLeu-----ProSerLysGln----- 149
 Db 3979 TTCTGCATCGCTCTCTCTGATGAGCCAGCAGAGCCCAAGGATGTCCCGAAGCTTT 4038
 Qy 150 -----Leu1yGluAerG1yVal1ThrVal1PheAlaVala1G1yValArg 163
 Db 4039 GTCCGCTACGTCCAGGGCTGTAGAGAAAGATCATGTGATCCGCGTGGCATTT--- 4095
 Qy 164 PheProArgTrg1uArg1uLeuH1a1aLeu1a1aSerGluProArg1yGlnH1aValLeu 183
 Db 4096 -----GGGCCCCATGGCCAACTCAAGCAGATCCG-----CTC 4128
 Qy 184 LeuAlaGluGlnVala1GluAerAla1ThraenG1yLeuPheSerThrLeuSerSerAla 203
 Db 4129 ATTGAGAAAGCAGGCCCTGAG-----AACAGGCC 4158
 Qy 204 IleCySerSerAla1ThrProAerCyuaerVala1GluAlaH1aProCyGluH1aArgThr 223
 Db 4159 TTCTGCTGACCGGCTGGATGAG-----CTGAGACAGCAAGG 4197
 Qy 224 LeuGluMetA1aTarg1uPheAlaG1yAen1aProCyuTrArg1ySerArgArgThr 243
 Db 4198 GAAGAAGATCGTTAGCTACCTGTGACCTTGCCCT-----GAAAGCCCTCCCTACT 4251
 Qy 244 LeuAlaVala1leuAla1aH1aCyuaerProPheTy1eTrblyAerVala1PheLeuThH1a 263
 Db 4252 CTGCC----- 4257
 Qy 264 ProAlaThrCy1yTArgThrThrCyuaerProG1yProCyuaerSerGln-----Pro 280
 Db 4258 CCCACATGACCAAGTACTGTGGGCCCCGGGCTTTGGGAGCTTTCGACCCCTGGGGGCC 4317
 Qy 281 Cyuaer1aenG1yG1yThrCyua-----ValProGluG1yLeuAerG1yTy1eGln 296
 Db 4318 AAGAGAACTCCATGCTGTGATGTGGCGTTTGCCCGGAAGATTGGAACAATTT--- 4374
 Qy 297 CyuaerCyuaerProLeuAlaPheG1yG1yGluAlaenCyua1aLeuLysLeuSerLeuGlu 316
 Db 4375 -----GTTGAAGCCGAG----- 4386
 Qy 317 CyuaerVala1aerLeuLeuPheLeuLeuAerSerSerAlaG1yThrThrLeuAerG1yPhe 336
 Db 4387 -----TTT 4389
 Qy 337 LeuArgAla1yVala1PheVal1yAerPheVal1---ArgAlaVala1LeuSerGluAerSer 355
 Db 4390 AACAGAGCAAGAGATTGATGAGAGGATTCACGCGATGATGTGGCCAGAGACAC 4449
 Qy 356 ArgAlaAerG1yG1yVala1a1aThrTy1eSerArg1yGluLeuLeuVala1a1aVala1ProValG1y 375
 Db 4450 ---ATCCACGTCAAGGTGCTGACATTCCTTACATGATGACCGTGGAGTACCCCTTGACG 4506

Qy 376 G1uTy1eGlnAerVala1ProAerLeuVala1TrpSerLeuAerG1y1leProPheArgG1yG1y 395
 Db 4507 GAGGACAGCTCCAAAGGGACATCTTCAGCGGGGTGGAGAGATCCGCTACAGAGCGGC 4566
 Qy 396 ---ProThrLeuThrG1ySerAlaLeuAerGlnAla1aG1uArgG1yPheG1ySerAla 414
 Db 4567 AACAGACCAACACTGGGCTGGCGCTGGCTGACTTGTGACCAAGCTTC---TTGGTC 4623
 Qy 415 ThrArgThrG1yGlnAerAerProAerAerVala1Vala1LeuThrGlnG1ySerH1aSer 434
 Db 4624 AGCCAGGGTACCGGAGACAGCGGCCCACTGATGTACATGATGTACCGGAATCTGGCC 4683
 Qy 435 G1uAerG1yVala1aG1yProAlaArgH1aAlaArgH1aAerG1yLeuLeuLeuG1y 454
 Db 4684 TCGATGAGATC-----AAGAGGCTCTCGAGACATCCAGAGTGGTGGCCATTGGA 4734
 Qy 455 ValG1ySerGluAlaVala1ArgAlaG1yLeuGlnGlu1leThrG1ySerProLysVala1 474
 Db 4735 GTGGGCTTAATGCCAACTGACAGAGCTGAGAGGATTT---GGCTGGCCCAATGCCCCC 4791
 Qy 475 MetValTy1eSerAerProGlnAerLeuPheAenGln1leProGluLeu---GlnG1yLys 493
 Db 4792 ATCTCATCCAGACTTTTGACAGCTCCCGGAGAGCTCTGACCTGTGGCTGACAGAG 4851
 Qy 494 LeuCySerSerArgGln-----ArgProG1yCyuaerThr 504
 Db 4852 TGCTGCTCCGGAAGAGGGGCTGCAGATCCCACTCTCCCTGCCCTGACTGC---AGC 4908
 Qy 505 GlnAlaLeuAerLeuVala1PheMetLeuAerThrSerAlaSerVala1G1yProGluAerPhe 524
 Db 4909 CAGCCCTGAGACGTATCTCTCTCTGAGAGGCTCTCCAGTTTCCAGCTTCTTAATTT 4968
 Qy 525 AlaGluMetGlnSerPheVala1ArgSerCyua1leuGlnPheGluVala1aenProAerPua1 544
 Db 4969 GATGAATGAAGAATTGTGTAAGCTTTTATTTCAAAAGCCAAATATAGGCTCATCTC 5028
 Qy 545 ThrGlnVala1G1yLeuVala1Ty1yG1ySerGlnVala1GlnThra1aPheG1yLeuAerThr 564
 Db 5029 ACTCAGGGGTCAGTGTCCAGTATGAGCATACACACCATTCGACGTCGATGAGACGTG 5088
 Qy 565 LysProThrArgAla1aMetLeuAerGluAla1leSerGlnAlaProTy1eLeuG1yVala1 584
 Db 5089 GCCCCGGAAGAACCCCTTTGTGCTGAGCTTGTGACCTGTGACCTGACGCGGAGAGGCCCC 5148
 Qy 585 G1ySerAlaG1yThra1aLeuLeuH1a1eTy1aerPua1MetThra1aGlnArgG1y 604
 Db 5149 AGCCAAATCCGAGATGCTGGGCTTGTGCTGCGATCTTGACTTCAGAAATGCATGCT 5208
 Qy 605 AlaArgProG1yVala1ProLysAlaVala1Vala1LeuThrG1yArgG1yAlaGluAer 624
 Db 5209 GCCAGGCGGAGGCTCAAAAGCGGTGTATCTGTGTACCGGACGTCTGTGATTTCA 5268
 Qy 625 AlaAlaVala1ProAlaGlnLysLeuAerAenG1y1leSerVala1LeuVala1G1yVala1 644
 Db 5269 GTGATGACACAGCTGATGCGGCCAGGCTCCAAAGATGACAGTTCCTGATTTGAAT 5328
 Qy 645 G1yProVala1LeuSerGluG1y---LeuAerArgLeuAlaG1yProArg---AerSer 661
 Db 5329 GAGATGCTTACATGACAGCCCAAGCTACGATCTTGCAAGGCCACAGAGCGACTCCAAAC 5388
 Qy 662 Leu1leH1aVala1a1aTy1a1aerPua1ArgTy1eH1a----- 674
 Db 5389 GTGTGAAGCTCCAGCAATGCAAGACTCC---TACATGTGTACCTTGGCAATTCCTT 5447
 Qy 675 -----GlnAerVala1leu1leGluTrbLeuCyua----- 683
 Db 5448 CTTCACAAACTGTGCTC-----TGATTTGTTAGATTTCGATGATGAGATGAGATGAGAA 5501
 Qy 684 ---G1yGluAla1yAerGlnProVala1aenLeuCyua1yAerProSerPro-----Cyua 699
 Db 5502 TGAAGAAGGCGCGGAGACTGTGACCTTGCCAGACCAAGTGCACACCGTGACTTGCA 5561